MEDICINSKE NAUKE ZASNOVANE NA DOKAZIMA

Slobodan Janković

Kragujevac, 2026

Autor: prof. dr Slobodan Janković, redovni professor Fakulteta medicinskiih nauka, Univerziteta u Kragujevcu

Recenzenti:

prof. dr Marina Kostić, redovni professor Fakulteta medicinskiih nauka, Univerziteta u Kragujevcu

prof. dr Jasmina Milovanović, redovni professor Fakulteta medicinskiih nauka, Univerziteta u Kragujevcu

Tehnički urednik: Slobodan Janković

Lektor i korektor: Nera Janković, filolog romanista

Izdavač: Fakultet medicinskih nauka, Univerzitet u Kragujevcu

Za izdavača: prof. dr Vladimir Janjić, dekan Fakulteta medicinskih nauka

Tiraž: 300

Štampa: Interprint, Kragujevac

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

616-07(075.8)

ЈАНКОВИЋ, Слободан, 1961-
Medicinske nauke zasnovane na dokazima / Slobodan Janković. - Kragujevac : Univerzitet, Fakultet medicinskih nauka, 2026 (Kragujevac : Interprint). - 51 str. : graf. prikazi, tabele ; 24 cm

Tiraž 300. - Bibliografija: str. 48-51.

ISBN 978-86-7760-281-9

a) Болести -- Дијагностика

COBISS.SR-ID 191913481

Odlukom izdavačke komisije br. 07-2643/4 od 30.03.2026. ova knjiga se štampa kao udžbenik za osnovne studije.

SADRŽAJ

UVOD 4

KAKO SPROVODITI MZD U PRAKSI? 5

PRETRAŽIVANJE MEDLAJN BAZE PODATAKA 7

POSTUPAK SA PRONAĐENIM PUBLIKACIJAMA RELEVANTNIM ZA PITANJE KOJE SMO POSTAVILI 7

KRITERIJUMI ZA ANALIZU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJE NEKOG TERAPIJSKOG METODA (NPR. LEKA) .. 8

KRITERIJUMI ZA ANALIZU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJA DIJAGNOSTIČKIH POSTUPAKA 10

KRITERIJUMI ZA PROCENU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJE PROGNOSTIČKOG FAKTORA 13

VALIDNOST I KLINIČKI ZNAČAJ STUDIJA KOJE SE BAVE UTVRĐIVANJEM NEŽELJENIH DEJSTAVA LEKOVA 14

SISTEMATSKI PREGLEDNI ČLANAK 15

VODIČI DOBRE PRAKSE 19

FARMAKOEKONOMSKE STUDIJE 21

VRSTE KLINIČKIH STUDIJA 24

DA LI JE U OBJAVLJENOJ STUDIJI KORIŠĆEN ADEKVATAN STATISTIČKI TEST ZA OBRADU REZULTATA? 26

NAJČEŠĆE GREŠKE U STATISTIČKOJ ANALIZI REZULTATA 30

METODOLOGIJA KRITIČKE PROCENE OBJAVLJENIH KLINIČKIH STUDIJA 32

SISTEMATSKA GREŠKA U SPROVOĐENJU KLINIČKIH STIDIJA ILI PRISTRASNOST (ENGL. “BIAS”) 38

NEDOSTAJUĆI PODACI 42

PROCENA VELIČINE EFEKTA U KLINIČKIM STUDIJAMA 43

ISHODI KLINIČKIH STUDIJA 45

STUDIJE DOKAZIVANJA EKVIVALENTNOSTI ILI NE-INFERIORNOSTI NOVE TERAPIJE U ODNOSU NA STANDARDNU 47

ALTERNATIVA MEDICINSKIM NAUKAMA ZASNOVANIM NA DOKAZIMA 48
LITERATURA 48

UVOD

Medicinske nauke zasnovane na dokazima (medicina, stomatologija, farmacija - MZD) podrazumevaju pažljivu, promišljenu upotreba najboljih dokaza iz biomedicinske literature u donošenju odluke o lečenju konkretnog bolesnika. To znači da se individualno kliničko iskustvo mora integrisati sa rezultatima (dokazima) kliničkih studija, koje su metodološki bile dobro postavljene, i čiji rezultati imaju klinički značaj, tj. njihova primena u praksi ima smisla. Sa današnjom dostupnošću medicinske literature putem Interneta, lečenje bolesnika samo na osnovu ličnog iskustva i znanja stečenog iz udžbenika, uz ignoraciju rezultata najnovijih, validno izvedenih studija, više nije ni stručno, ni moralno prihvatljivo. Samo zbog naše neinformisanosti, bolesnik može ne dobiti najefikasniju i najbezbedniju terapiju, čime su značajno smanjene njegove šanse za izlečenje. Bolesnik takođe treba da učestvuje u donošenju svih važnih odluka koje se tiču njegovog tretmana; i lekar, stomatolog i farmaceut su dužni da bolesniku objasne sve raspoložive mogućnosti, jezikom koji on razume, i potrude se da svaka odluka bude donesena sporazumno.

MZD zasnovana na dokazima je integracija kliničkog iskustva, sistema vrednosti pacijenta i najboljih dokaza iz literature. Dokazi sami po sebi ne dovode do zaključaka, već samo pomažu u određivanju najboljeg tretmana pacijenta. Primena MZD obično se započinje prilikom pregleda pacijenta, kada nastaju pitanja o efektima lečenja, neželjenim dejstvima, korisnosti dijagnostičkog testa, prognozi bolesti, troškovima lečenja i dr. Na najveći broj pitanja lekar, stomatolog i farmaceut nisu u mogućnosti da odgovor daju odmah, zbog njihove složenosti (statistika je pokazala da se na

svaka 3 bolesnika pojave 2 pitanja na koje se odgovor ne može dati odmah), već moraju potražiti rešenje u medicinskoj literaturi. MZD zahteva nove veštine od lekara i farmaceuta: pretraživanje literature i primenu određenih pravila za procenu kvaliteta dokaza iz literature.

U svetu postoji više hiljada medicinskih časopisa u kojima se godišnje objavi nekoliko desetina hiljada radova. Nažalost, istraživanja su pokazala da je svega 1% objavljenih studija metodološki bilo dobro urađeno, tako da se njihovi rezultati mogu bezrezervno primeniti u praksi. Da bi se snašli u takvoj prašumi publikacija, i izdvojili samo validne studije, neophodno je dobro poznavati i primenjivati kriterijume za procenu kvaliteta metodologije objavljenih studija.

Ideju medicinskih nauka zasnovanih na dokazima prvi je formulisao i promovisao američko-kanadski epidemiolog Dejvid Saket. On je još 1967. godine osnovao prvu katedru kliničke epidemiologije i statistike na MekMaster univerzitetu u Kanadi, a zatim je osamdesetih godina XX veka razvio metodologiju kritičke procene medicinske literature. Tokom 1994. godine prešao je u Oksford, gde je osnovao Centar za medicinu zasnovanu na dokazima, i 1996. izdao prvu knjigu "Medicina zasnovana na dokazima" zahvaljujući kojoj se filozofija kritičke procene medicinske literature i primene najboljih saznanja iz literature u kliničkoj praksi proširila svetom.

KAKO SPROVODITI MZD U PRAKSI?

U radu sa bolesnicima se spontano javljaju problemi koje moramo rešavati, i na osnovu kojih mi postavljamo pitanja, na koja potom tražimo odgovor. Problemi se javljaju u postavljanju dijagnoze bolesti, u otkrivanju njenog uzroka, u izboru dijagnostičkih metoda koje će najpre dovesti do dijagnoze, u izboru načina lečenja i u određivanju prognoze bolesnika. Rešenje problema mnogo zavisi od pravilnog postavljanja pitanja; ako se ono ne postavi dobro, čitava potraga za odgovorom će skrenuti na pogrešan put. Dobro postavljeno pitanje uvek ima 4 elementa: precizno definisano

stanje bolesnika, intervenciju koja se planira, alternativnu intervenciju sa kojom se planirana intervencija upoređuje, i ishod koji se očekuje. Evo primera dobro postavljenog pitanja: «Da li kod bolesnika sa hroničnom srčanom insuficijencijom 3. stepena (stanje bolesnika) primena ACE-inhibitora (planirana intervencija) smanjuje smrtnost bolesnika (ishod) više ili manje od blokatora receptora za angiotenzin (alternativna intervencija)». Ova četili elementa dopro postavljenog istraživačkog pitanja se pamte po mnemotehničkoj formuli „PICO“:

P – pacijent

I – intervencija, tj. lečenje koje se ispituje

C – komparator, od engleske reči „Comparator“

O – ishod lečenja koji se ispituje, od engleske reči „Outcome“

Gde i kako tražimo odgovore na pitanja koja postavljamo tokom tretmana bolesnika? Pre svega treba pretražiti baze podataka na Internetu koje sadrže sistematske pregledne članke o pojedinim klinčkim problemima. Najpouzdanija od takvih baza je Kohranova baza podataka (Cochrane database of systematic reviews) jer su svi sitematski pregledni članci u njoj rađeni prema metodološkim principima MZD, dakle validni su. Ova baza podataka je besplatna za korišćenje sažetaka sistematskih pregleda, koji su strukturirani i dovoljno obimni, tako da pun tekst pregleda često nije neophodan. Ako odgovor na pitanje koje smo postavili ne nađemo u Kohranovoj bazi, preostaje nam da sami prikupimo relevantnu literaturu i procenimo njenu validnost i kliničku značajnost. Literaturu ćemo najefikasnije prikupiti pretraživanjem MEDLAJN (MEDLINE) baze sažetaka radova publikovanih u više hiljada medicinskih časopisa, koja je dostupna putem Interneta i besplatna za korišćenje. Druga mogućnost je da se pretražuje GOOGLE SCHOLAR, takođe besplatna baza naučnih časopisa. Osim te dve opšte baze, postoje i druge, koje zahtevaju pretplatu: SCOPUS, EBSCO, EMBASE. Takođe postoje i baze podataka specijalizovane za pojedine vrste publikacija, kao što su farmakoekonomske studije i slično.

PRETRAŽIVANJE MEDLAJN BAZE PODATAKA

MEDLAJN baza sažetaka publikovanih studija se pretražuje pomoću ključnih reči. Kada se u prozor za pretraživanje unesu ključne reči, i pokrene pretraživanje tasterom „GO“, na ekranu će se pojaviti sažeci svih radova u kojima su pomenute te ključne reči. Prema tome, veština pretraživanja se sastoji u dobrom izboru ključnih reči, i njihovom mudrom kombinovanju. Ako su ključne reči isuviše opšteg značenja, ili ako koristimo samo jednu ključnu reč, pretraživač će nam „izbaciti“ više stotina, pa i hiljada sažetaka, koje nećemo biti u mogućnosti da sve pročitamo. Ako pak koristimo ključne reči suviše uskog značenja, ili koristimo veći broj reči (pet, šest ili više), pretraživač će propustiti veliki broj relevantnih radova, i na ekranu će se pojaviti samo nekoliko sažetaka.

Korisno je birati ključne reči iz rečnika koji koristi sama baza podataka (tezaurusa), jer će se tada propustiti manje relevantnih radova. Tezaurus se nalazi na linku “MESH” Medlajn baze podataka. Takođe se mogu pri pretraživanju aktivirati razni ograničavajući faktori (tzv. filteri), koji usmeravaju pretraživanje na samo određene delove baze podataka. Na primer, ako aktiviramo ograničavajući faktor „Clinical trial“, pretraživanje će se ograničiti samo na publikacije koje opisuju rezultate kliničkih studija, i neće nam se pojavljivati sažeci studija rađenih na životinjama, niti sažeci preglednih članaka.

Principi pretraživanja drugih baza podataka, koje su prethodno pomenute, su isti kao i principi pretraživanja MEDLAJN baze. Može se desiti da neke od baza imaju manje opcija za pretraživanje ili nemaju definisan tezaurus, ali to ne menja značajno postupak pretraživanja.

POSTUPAK SA PRONAĐENIM PUBLIKACIJAMA RELEVANTNIM ZA PITANJE KOJE SMO POSTAVILI

Kada pronađemo sažetke publikacija koje su relevantne za naše pitanje, treba pronaći te publikacije u celini, pročitati ih i analizirati njihovu

metodološku ispravnost (validnost) i kliničku značajnost. I validnost i klinička značajnost se utvrđuju po kriterijumima koji su različiti za različite tipove studija: dijagnostičke, terapijske, studije prognoze, itd. Do sada formulisani kriterijumi, koji se pominju u većini udžbenika Medicine zasnovane na dokazima, biće navedeni u daljem tekstu, po vrstama studija koje se analiziraju. Međutim, to ne znači da su navedeni kriterijumi i jedini koji se mogu koristiti. Pravilan izbor testa za statističku obradu rezultata takođe može biti važan kriterijum za procenu validnosti neke studije, i o tome će biti reči u posebnim poglavljima. Takođe, čitalac može otkriti neke druge metodološke neispravnosti studije koju analizira, na osnovu kojih će zaključiti da studija nije validna. Dakle, kriterijumi koji budu pomenuti u daljem tekstu su osnova analize validnosti i kliničke značajnosti, i mogu se dopunjavati i proširivati prema sposobnosti analitičara.

Najzad, kada smo pronašli relevantne publikacije za kliničko pitanje koje smo postavili, i kada smo proverili njihovu validnost i klinički značaj, na osnovu sumacije rezultata validnih i značajnih publikacija formiramo plan lečenja i /ili dopunske dijagnostike pacijenta. Time postižemo da naš pacijent dobije najbolje rešenje za svoj zdravstveni problem koje savremena medicinska nauka može pomuditi.

KRITERIJUMI ZA ANALIZU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJE NEKOG TERAPIJSKOG METODA (NPR. LEKA)

Studije koje procenjuju efikasnost i bezbednost nekog terapijskog postupka ili leka na pacijentima se nazivaju kliničkim studijama u užem smislu reči. Na njima se zasniva savremena terapijska praksa, tako da je od ogromnog značaja pravilno proceniti validnost i kliničku značajnost takvih studija, i odbaciti one koje nas zavode na pogrešan put. Validna klinička studija mora da ispuni sledeće uslove: mora biti randomizirana (tj. pacijenti su po principu slučajnog izbora svrstani u eksperimentalnu i kontrolnu grupu) i dvostruko slepa (tj. ni ispitivači, ni pacijenti nisu znali ko prima placebo, a ko aktivnu kontrolu); i kontrolna grupa i grupa koja prima novi

tretman moraju biti po svemu slične na početku studije, i moraju se tretirati na isti način; svi pacijenti koji su ušli u studiju moraju biti uzeti u obzir prilikom izvođenja zaključaka, i analizirani u grupama u kojima su bili na početku (dozvoljivo je da najviše 20% bolesnika koji su ušli u studiju istu ne završi do kraja, već ranije «ispadne» iz daljeg lečenja i praćenja). Ukoliko neki od ovih kriterijuma nije ispunjen, mi ne možemo biti sigurni da je rezultat studije posledica stvarnog efekta novog terapijskog postupka ili leka, a ne nekog nekontrolisanog faktora koji se slučajno ili namerno umešao u lečenje bolesnika. Da bi se olakšala provera validnosti neke kliničke studije, međunarodna stručna tela su donela vodič za sastav protokola kliničke studije, kao i vodič za izveštavanje o sprovedenoj kliničkoj studiji u publikaciji objavljenoj u naučnom časopisu. Vodič za sastav protokola se naziva skraćeno SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), a vodič za izveštavanje o sprovedenoj kliničkoj studiji CONSORT (CONsolidated Standards Of Reporting Trials). Detaljni vodiči i liste za podsećanje se mogu naći na veb stranici https://www.consort-spirit.org/ .

Značajnost kliničke studije se može proceniti samo ako se efekat pokazan u studiji stavi u kontekst stvarnog života, uzimajući u obzir stvarnu prevalenciju oboljenja koje se tretira. Ako je primećeno, na primer, da novi lek smanjuje rizik od smrtnog ishoda za polovinu (tj. relativno smanjenje rizika je 50%), to još uvek ne znači mnogo, dok ne ustanovimo koliko je smrtni ishod bio učestao bez novog leka. Ako je bez novog leka umiralo oko 2% bolesnika, a sa novim lekom 1% bolesnika, tačno je da je rizik smanjen relativno za 50% (relativno smanjenje rizika = 1%/2% = 0.5 = 50%), ali je apsolutno smanjenje rizika zapravo samo 1%! (apsolutno smanjenje rizika = 2% - 1% = 1%). Šta to u stvari znači možemo shvatiti ako zamislimo sledeće: apsolutno smanjenje rizika za 1% zapravo znači da će od 100 pacijenata tretiranih novim lekom umreti 1 manje no inače, ili, drugačije rečeno, da treba da lečimo 100 pacijenata novim lekom, da bi kod jednog od njih izbegli smrtni ishod! Sada nam efekat hipotetičkog leka izgleda veoma mali, odnosno beznačajan. Da je učestalost smrtnog ishoda kod pacijenata bez novog leka bila, recimo, 10%, prepoloviti rizik novim lekom (relativno smanjiti rizik za 50%) bi značilo smanjenje učestalosti smrtnog ishoda na 5%. Tada bi apsolutno smanjenje rizika bilo: 10% - 5% = 5%. To bi značilo da će od 100 pacijenata lečenih novim lekom umreti 5 manje no inače, tj. da

treba da lečimo 100 pacijenata da bi kod petorice izbegli smrtni ishod, ili, što je isto, da lečimo 20 pacijenata da bi kod jednog izbegli smrtni ishod (jer je 100/5 = 20)! U ovoj drugoj hipotetičkoj situaciji efekat leka je znatno veći, i približava se značajnosti.

Dakle, kada analiziramo značajnost kliničke studije, moramo uzeti u obzir apsolutno smanjenje rizika koje nosi novi lek, i utvrditi broj bolesnika koje treba lečiti novim lekom, da bi se izbegao neželjeni ishod kod jednog od njih (engleski: "Number Needed to Treat - NNT). Smatra se da značajan efekat postoji kada treba lečiti manje od 15 bolesnika da bi se izbegao neželjeni ishod kod jednog od njih, ali to je samo slobodna procena, koja dosta zavisi od težine neželjenog ishoda. Ako je neželjeni ishod smrt, onda možda vredi lečiti i 20 bolesnika da bi se jedan od njih spasao.

Ako smo utvrdili da je klinička studija koju smo analizirali validna i klinički značajna, postavlja se pitanje da li se rezultati validne i značajne studije mogu primeniti kod našeg pacijenta? Verovatno mogu, ako pacijent nije isuviše različit od onih u studiji, i ako će novi tretman i njegove posledice zadovoljiti vrednosti i želje našeg pacijenta. Dakle, odgovor na ova pitanja se može doneti samo u razgovoru sa pacijentom.

KRITERIJUMI ZA ANALIZU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJA DIJAGNOSTIČKIH POSTUPAKA

Studija nekog dijagnostičkog postupka (testa) se može proceniti kao metodološki ispravna, ako se na sledeća pitanja može odgovoriti sa «DA»:

- Da li su rezultati dijagnostičkog postupka o kome je reč u studiji nezavisno i slepo upoređeni sa «ZLATNIM STANDARDOM» dijagnoze oboljenja koje se dokazuje tim dijagnostičkim postupkom? Na primer, da li je povišen nivo nekog serumskog markera tumora jetre kod bolesnika upoređen sa rezultatom biopsije jetre tog istog bolesnika, koja treba da potvrdi ili opovrgne prisustvo tumora.

• Da li je dijagnostički test ispitan na odgovarajućem spektru pacijenata (sličnog onom u stvarnom životu)? Rezultati studije o dijagnostičkom testu su korisni ako je dijagnostički test primenjen na bolesnicima sa blagom, ali i teškom bolešću, na bolesnicima u ranoj i kasnoj fazi, na nelečenim i lečenim bolesnicima, kao i na zdravim osobama. Takođe, korisno je da test bude primenjen i na bolesnicima sa bolestima koje su slične onoj koja se ispituje.

• Da li je zlatni standard upotrebljen nezavisno od rezultata dijagnostičkog testa? Istraživači često žele da dokažu da je novi test koristan, pa mogu da požele i da potvrde pozitivne rezultate novog testa pozitivnim rezultatima „zlatnog standarda“. Zato je neophodno da se obezbedi nezavisno upoređenje novog testa sa „zlatnim standardom“.

Za objavljivanje studija dijagnostičkih testova takođe postoji međunarodni vodič, koji reguliše šta sve treba navesti iz metodologije i rezultata. Vodič se naziva STARD (An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies), i može se naći na veb stranici https://www.equator-network.org/reporting-guidelines/stard/ .

Novi dijagnostički postupak može biti klinički značajan ako je posle njegove primene verovatnoća da pacijent ima određenu bolest značajno veća od verovatnoće pre primene tog testa. Da bi se ta promena u verovatnoći utvrdila, potrebno je prethodno izračunati nekoliko parametara: senzitivnost, specifičnost, prediktivnu vrednost testa i odnos mogućnosti. Način izračunavanja ovih parametara se može videti na primeru studije koja se bavila kliničkom značajnošću kompjuterizovane tomografije (CT) u dijagnostici disekcije aorte. U sledećoj tabeli navedeni su rezultati CT-pregleda kod pacijenata sa sumnjom na disekciju aorte.

Tabela 1. Primer izračunavanja parametara dijagnostičke vrednosti testa.

CT		Akutna disekcija aorte		Ukupno
		Prisutna	Odsutna	
Rezultat CT-pregleda	Pozitivan	a (**45**)	b (**4**)	**a+b** (49)
	Negativan	c (**3**)	d (**27**)	**c+d** (30)
Ukupno		**a+c** (48)	**b+d** (31)	**a+b+c+d** (79)

- Senzitivnost (procenat bolesnika kod kojih je test pozitivan) se izračunava na sledeći način: a/(a+c) = 45/48 = 0,94 = 94 %
- Specifičnost (procenat zdravih osoba kod kojih je test negativan): d/(b+d) = 27/31 = 0,87 = 87 %
- Odnos mogućnosti za pozitivan ishod testa (tj. sposobnost razlikovanja stvarno pozitivnog rezultata od lažno pozitivnog rezultata) = senzitivnost/(100-specifičnost) = 94/(100-87) = 7,2
- Odnos mogućnosti za negativan ishod testa (tj. sposobnost razlikovanja lažno negativnog od stvarno negativnog rezultata) = (100-senzitivnost)/specifičnost = (100-94)/87 = 0,07
- Pozitivna vrednost predviđanja (tj. u kojoj meri je pozitivan rezultat zaista tačan) = a/(a+b) = 45/49 = 92 %
- Negativna vrednost predviđanja (tj. u kojoj meri je negativan rezultat testa zaista tačan) = d/(c+d) = 27/30 = 90%
- Prevalencija oboljenja u grupi pacijenata sa sumnjom na disekciju aorte se može grubo proceniti na osnovu podataka iz same studije = (a+c)/(a+b+c+d) = 48/79 = 0,61= 61%
- Odnos pre testa (engleski pre-test odds) = prevalencija/(1-prevalencija) = 61/39 = 1,56
- Odnos posle testa (engleski post-test odds) = odnos pre testa * odnos mogućnosti = 1,56 * 7,2 = 11,2
- Verovatnoća posle testa = odnos posle testa/(odnos posle testa + 1) = 11,2/12,2 = 0,92 (92 %).

Dakle, verovatnoća da će pozitivan nalaz disekcije aorte na CT-pregledu značiti da bolesnik ZAISTA ima disekciju aorte je 92%. Ako je prethodno verovatnoća da bolesnik ima disekciju aorte na osnovu samo anamneze i fizikalnog pregleda bila, recimo, 50%, onda je primena CT-a klinički opravdana, jer značajno povećava verovatnoću dijagnoze.

Ipak, ovako procenjena klinička značajnost još uvek ne znači da ćemo test (CT-pregled) primeniti kod konkretnog bolesnika. Ostaju pitanja: da li je test dostupan, da li će biti precizan kod našeg bolesnika, da li će rezultat testa nešto promeniti u lečenju bolesnika i da li će bolesnik prihvatiti da se podvrgne testu. Na ova pitanja lekar može da odgovori samo ako ostvari partnerski odnos sa bolesnikom, tako da sve odluke donosi u dogovoru sa njim.

KRITERIJUMI ZA PROCENU VALIDNOSTI I KLINIČKE ZNAČAJNOSTI STUDIJE PROGNOSTIČKOG FAKTORA

Prognostički faktor je svaka osobenost pacijenta koja može imati uticaj na pojavu oboljenja (ako pacijent prethodno nije imao znake bolesti) ili na ishod oboljenja (ako je pacijent već oboleo). Na primer, svi znamo da muške osobe koje imaju obim struka veći od 104 cm dva puta češće dožive infarkt miokarda od osoba koje nisu gojazne. U ovom slučaju prognostički faktor je obim struka.

Da bi studija prognostičkog faktora bila validna, ona mora da prati reprezentativan uzorak pacijenata, koji je dovoljno veliki, i dovoljno dugo da se vidi kakav će biti ishod bolesti. Na primer, ako se ispituje uticaj starosti na preživljavanje upale pluća, dovoljno je pacijente pratiti tridesetak dana, jer će u tom periodu doći ili do ozdravljenja, ili do smrti. Ali, ako pratimo uticaj svakodnevnog uzimanja aspirina na učestalost gastrointestinalnih karcinoma, takve pacijente moramo pratiti godinama, pa i decenijama. Kod validnih studija istraživači uspeju da ne „izgube“ pacijente, tj. da isprate do kraja najveći broj pacijenata koji su ušli u studiju; dozvoljivo je da se „izgubi“ najviše 20% pacijenata.

Takođe je važno da se ishod oboljenja utvrdi objektivno, tj. da neki drugi stručnjak, a ne sami istraživači utvrdi šta se dogodilo sa pacijentom, ne znajući da li on ima prognostički faktor ili ne. U protivnom, često istraživači žele da „dokažu“ da prognostički faktor zaista utiče na ishod, pa mogu videti ishod i tamo gde ga nema.

Vodič za adekvatno izveštavanje o sprovedenoj prognostičkoj studiji je isti kao i za studije dijagnostičkih testova: već pomenuti STARD.

Prognostički faktor će imati klinički značaj ukoliko daje dovoljno preciznu prognozu, tj. ako interval poverenja oko prognoze nije preveliki. Na primer, ako primena tetraciklina u prvom trimestru trudnoće može povećati učestalost urođenih anomalija na 4 ± 4 %, to nam ne znači mnogo, jer se učestalost anomalija dece rođene iz trudnoća u kojima majke nisu uzimale

nikakve lekove kreće oko 2 ± 2%, dakle postoji veliko preklapanje. Ako bi studija našla da je učestalost anomalija posle primene tetraciklina 4 ± 0.5%, tada bi već klinički značaj primene tetraciklina u trudnoći bio sasvim jasan. Takođe, za kliničku značajnost je bitna i učestalost ishoda koji se prati. Na primer, smrtni ishod kod običnih akni je tako redak, da nema smisla ispitivati uticaj bilo kog faktora na njega.

VALIDNOST I KLINIČKI ZNAČAJ STUDIJA KOJE SE BAVE UTVRĐIVANJEM NEŽELJENIH DEJSTAVA LEKOVA

Validnost ovih studija se najpre utvrđuje primenom istih kriterijuma koji se koriste kod studija efekata nekog metoda lečenja (videti ranije). Pored toga se ovde mora utvrditi i uzročnost neželjenog događaja, tj. njegova povezanost sa lekom. Uzročnost utvrđujemo na osnovu procene četiri elementa: (1) da li je primena leka prethodila neželjenom događaju ili ne (engl. “challenge”); (2) da li je posle primene leka došlo do povlačenja neželjenog događaja (engl. “dechallenge”); (3) da li je posle ponovnog uvođenja leka u terapiju došlo do ponovne pojave istog neželjenog događaja (engl. “rechallenge”) i (4) da li postoji neki drugi uzrok neželjenog događaja osim primene leka. Ako su pozitivni challenge, dechallenge i rechallenge, a nema drugog mogućeg uzroka neželjenog događaja osim leka, onda je verovatnoća da je lek uzrokovao neželjeni događaj veoma visoka, pa više ne govorimo o neželjenom događaju, već o **neželjenom dejstvu** leka. Klinički značaj se utvrđuje zavisno od vrste studije u metodološkom pogledu. Ako je studija bila tipa «slučaj-kontrola», onda se kao rezultat dobija odnos mogućnosti (engleski “odds ratio“), koji je značajan samo ako je jednak ili veći od 4. Ako je studija bila tipa kliničke studije ili kohortne studije, onda treba videti koliki je relativni rizik od nastanka neželjenog dejstva; ako je on veći od 3, onda rezultat studije ima praktičnog smisla. Kako se računaju «odnos mogućnosti» i relativni rizik, može se videti iz sledeće tabele.

Tabela 2. Primer izračunavanja odnosa šansi ("odds ratio") i relativnog rizika.

		Neželjeno dejstvo		
		Prisutno	Odsutno	Ukupno
Izloženi leku	Da	a	b	a+b
	Ne	c	d	c+d
	Ukupno	a+c	b+d	a+b+c+d

Relativni rizik: RR=[a/(a+b)]/[c/(c+d)] Odds ratio: OR=ad/bc

Pored odnosa mogućnosti i relativnog rizika, korisno je odrediti i apsolutno povećanje rizika. Na primer, ako je apsolutni rizik od pojave urođene anomalije mozga kod dece čije majke u trudnoći nisu koristile lekove 1%, a kod dece čije majke su uzimale antiepileptik valproat 2%, onda je apsolutno povećanje rizika od anomalija mozga sa primenom valproata jednako 2% - 1% = 1%. To zapravo znači da 100 pacijentkinja treba u trudnoći da uzima valproat, da bi se kod deteta samo jedne od njih pojavile anomalije mozga na rođenju (100/1 = 100). Ovakvu veličinu, koja predstavlja recipročnu vrednost apsolutnog povećanja rizika, nazivamo «broj bolesnika koje treba lečiti da bi se kod jednog javilo neželjeno dejstvo», ili engleski "Number Needed to Narm" (NNH). Ako izračunamo NNH, bolje ćemo razumeti kakvom se zapravo riziku izlažu pacijenti uzimajući određeni lek.

Međunarodni vodič za izveštavanje o sprovedenim opservacionim studijama (u koje spadaju i studije neželjenih dejstava lekova) se naziva STROBE (skraćenica od "The Strengthening the Reporting of Observational Studies in Epidemiology"), i može se naći na veb stranici https://www.equator-network.org/reporting-guidelines/strobe/ .

SISTEMATSKI PREGLEDNI ČLANAK

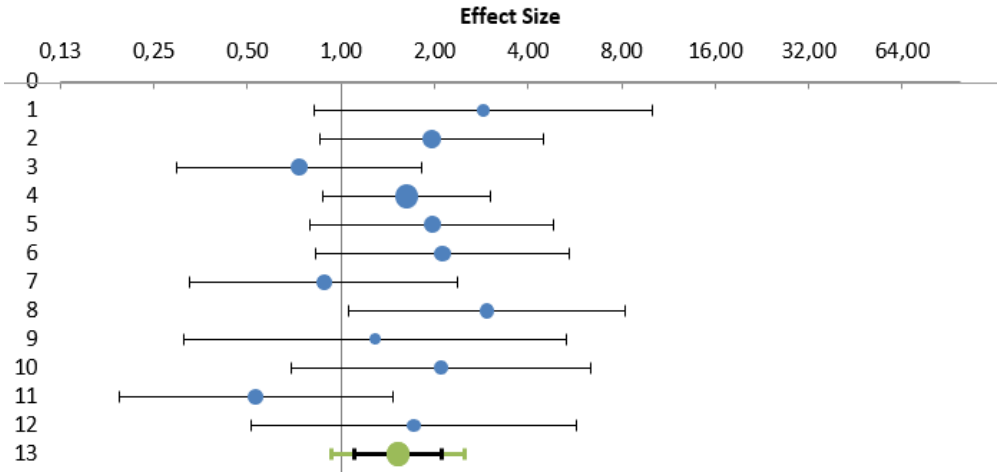
Pregledni članak sažima i tumači rezultate objavljenih kliničkih studija i meta analiza (meta analiza = pacijenti iz više pojedinačnih kliničkih studija istog leka se sakupe i zajedno statistički obrade, tako da izlazi kao da je rađena jedna veoma velika klinička studija), i pokušava da čitaocu pruži

sumarnu sliku nekog terapijskog ili dijagnostičkog pitanja. Osnovni uslov da pregledni članak pruži validne rezultate, jeste da bude urađen kao sistematski pregledni članak, tj. da njegovi autori sistematski pregledaju svu dostupnu literaturu (dakle da pretraže sve značajne baze publikovanih medicinskih časopisa), da je analiziraju na validnost i kliničku značajnost, i da u obzir uzmu samo validne studije. To znači da validan sistematski pregled mora imati odeljak pod nazivom «METOD», gde autori detaljno opisuju gde su sve tražili publikacije o temi o kojoj se piše, kako su pretraživali dostupne baze podataka i kako su ocenjivali metodološku ispravnost pronađenih publikacija. Kada pročitamo taj odeljak, treba da nam bude jasno šta je od pronađene literature otpalo, a šta je uzeto u obzir za izvođenje zaključaka.

Dobar sistematski pregled takođe procenjuje i kliničku značajnost rezultata koje obrađuje, i daje kritičko mišljenje o temi kojom se bavi.

Kada se čita sistematski pregledni članak koji u sebi sadrži i meta-analizu potrebno je poznavati osnovne metodološke principe njihove izrade kako bi se razumeli prikazani rezultati, kako brojčano, tako i u vidu grafika. Sumarni rezultat meta analize je obično prikazan na specifičnom grafiku koji se naziva “šuma grafik” ili engleski “forest plot”. Na slici niže prikazan je primer šuma grafika, gde su brojevima označene studije koje su ušle u meta analizu (njih 12), krugovi predstavljaju odnos šansi (engl. “odds ratio”) za postizanje efekta ako se koristi ispitivani lek umesto komparatora, a horizontalni graničnici označavaju granice poverenja za izračunate odnose šansi.

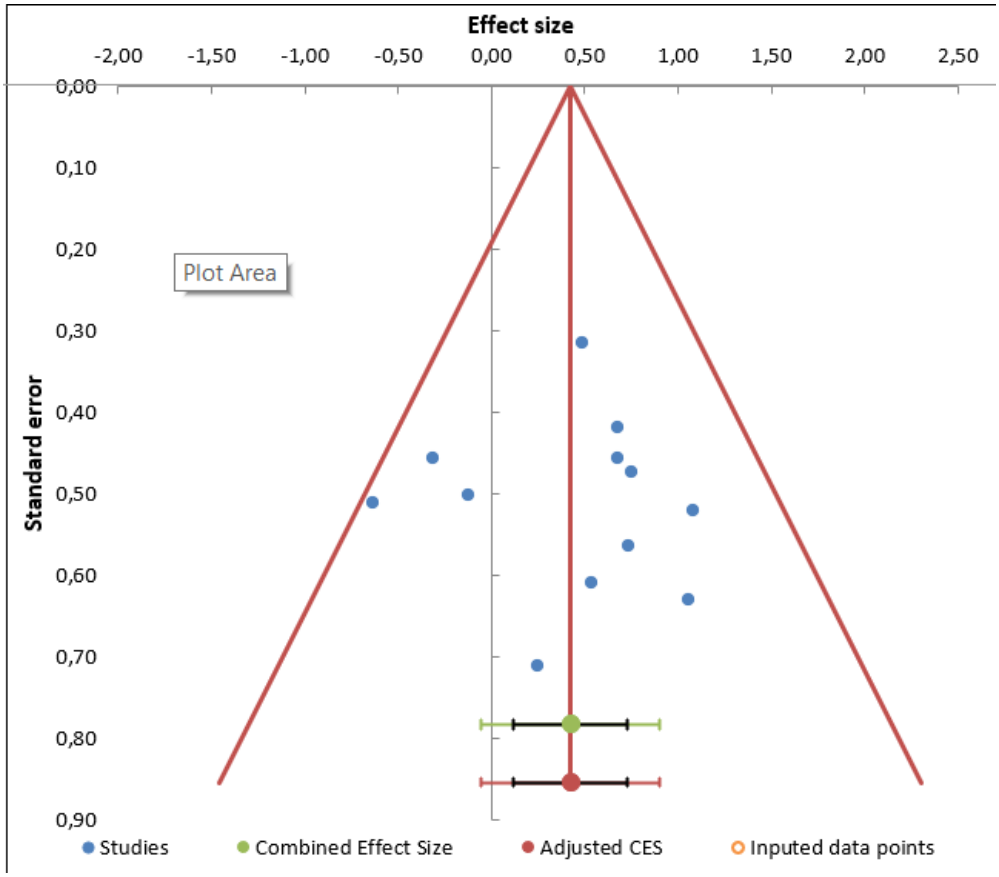
Slika 1. “Šuma grafik” (engl. “forest plot”). Na x-osi se nalazi odnos šansi da ispitivani lek postigne efekat u odnosu na komparator, a na y-osi redni broj studije koja je uključena u meta analizu.



Zeleni krug pod rednim brojem 13 označava sumarni odnos šansi, dobijen kao srednja vrednost odnosa šansi pojedinačnih studija koje su ponderisane varijabilnošću rezultata u svakoj studiji, tj. pomoću varijanse (što je veća varijansa, udeo odnosa šansi te studije u sumarnom odnosu šansi je manji, i obrnuto). Da bi sumarni rezultat bio statistički značajan, horizontalni graničnici oko sumarnog odnosa šansi ne smeju zahvatiti jedinicu.

Da li su sistematskim pregledom obuhvaćene sve relevantne studije u meta analizi se ispituje kroz konstrukciju tzv, "levak grafika" (engl. "funnel plot"). Na x-osi levak grafika se nalazi odnos šansi, a na y-osi varijabilnost u studijama (ovde izražena kao standardna greška). Kružići na grafiku predstavljaju pojedinačne kliničke studije. Ako su studije raspoređene simetrično oko srednje vrednosti efekta, onda verovatno nema studija koje nisu uzete u obzir. Ako pak postoji asimetrija, kao što je prikazano na grafiku niže, verovatno su neke studije propuštene, pa treba korigovati procenu srednje vrednosti efekta, odnosno odnosa šansi.

Slika 2. "Levak grafik" (engl. "funnel plot").



Na žalost, danas je malo sistematskih pregleda koji su metodološki pravilno sprovedeni. Veoma često su oni nepotpuni, ili neadekvatno analiziraju pronađenu literaturu, pa dolaze do pogrešnih zaključaka. To može biti veoma opasno, jer navodi čitaoce, lekare i farmaceute, na pogrešne zaključke, i na usvajanje neefikasnih ili štetnih metoda lečenja. Ponekada su pisci sistematskih pregleda pod uticajem farmaceutskih kompanija, pa zaključci njihovih radova nisu ništa više od trgovačke promocije leka neke kompanije. Zato pri čitanju sistematskih pregleda treba biti krajnje oprezan, i krajnje kritički proveriti da li ispunjavaju osnovne uslove validnosti. U pokušaju da se prepoznaju metodološki neispravno urađeni sistematski pregledi stvoren je instrument (upitnik) koji olakšava procenu kvaliteta sistematskih preglednih članaka. Naziv ovog instrumenta je AMSTAR 2 (A MeaSurement Tool to Assess Systematic Reviews 2), i sadrži 16 pitanja, o kojih su najvažnija:

- Da li je protokol sistematskog pregleda negde registrovan i objavljen pre početka njegovog sprovođenja
- Da li je način pretrage literature adekvatan
- Da li je procenjen rizik od subjektivnog uticaja na rezultate (engl. "bias")
- Da li je dato objašnjenje zašto su neke studije isključene iz sistematskog pregleda
- Da li su rezultati sumirani na adekvatan način (ako je u okviru sistematskog pregleda rađena i meta-analiza)

Pored vodiča za procenu validnosti sistematskih pregleda, postoji i vodič za izveštavanje o sprovedenom sistematskom pregledu pod nazivom PRISMA (dostupan je na adresi: https://www.equator-network.org/reporting-guidelines/prisma/).

VODIČI DOBRE PRAKSE

Vodiči dobre prakse su dokumenti koji između ostalog sadrže i preporuke čiji cilj je poboljšanje zdravstvene zaštite. Vodiče donose komisije sačinjene od eksperata iz oblasti kojoj je vodič namenjen, i njihove preporuke treba da budu zasnovane na najboljim dokazima iz kliničkih studija. Da bi se vodiču moglo verovati (tj. da bi bio validan) potrebno je da su zadovoljeni sledeći principi koje je postavio Medicinski institut iz Sjedinjenih Američkih Država:

- zasnovanost na sistematskom pregledu postojećih dokaza iz medicinske literature;
- vodič donosi stručna, multidisciplinarna komisija koja se sastoji od stručnjaka i predstavnika ključnih pogođenih grupa;
- vodič uzima u obzir važne podgrupe pacijenata i ono na šta bi pacijenti pristali, prema potrebi;

- proces stvaranja i prihvatanja vodiča mora biti eksplicitan i transparentan koji će sprečiti uticaj predrasuda stručnjaka i sukobe interesa;
- objašnjenje odnosa između terapijskih alternativa i ishoda lečenja treba da bude jasno, i za svaku preporuku treba da se oceni kvalitet dokaza i stepen preporuke;
- vodič mora da se povremeno ažurira, tj. preporuke moraju da se menjaju ako se u međuvremenu pojave novi dokazi sa drugačijim tvrdnjama.

Za svaku preporuku navedenu u vodiču neophodno je da komisija koja je pravila vodič odredi kvalitet (tj. sigurnost) dokaza iz literature i stepen preporuke za primenu u praksi. Kvalitet dokaza se može klasifikovati na više načina, ali se najčešće koristi tzv. GRADE klasifikacija koju je izradila međunarodna grupa eksperata (GRADE je skraćenica od „Grading of Recommendations Assessment, Development and Evaluation"). Po toj klasifikaciji, kvalitet dokaza može biti visok, osrednji, nizak ili vrlo nizak. **Visok kvalitet** dokaza postoji kada je velika verovatnoća da je stvarni efekat blizu procenjenog efekta u kliničkoj studiji. **Osrednji kvalitet** znači da je efekat u realnom životu verovatno blizak efektu dobijenom u kliničkoj studiji, ali postoji mogućnost da se značajno razlikuje. **Nizak kvalitet** se dodeljuje studijama čiji rezultat u pogledu efekta neke terapije može biti značajno drugačiji od efekta koji bi se mogao očekivati u realnom životu. Najzad, **vrlo nizak kvalitet dokaza** iz kliničkih studija znači da je stvarni efekat verovatno značajno drugačiji od efekta neke terapije dobijenog u kliničkoj studiji.

Stepen preporuke se obično klasifikuje kao A, B, C i D. **A je najviši stepen preporuke da koju treba obavezno primenjivati u praksi**, jer su pronađeni dobri dokazi da intervencija poboljšava važne zdravstvene ishode, pa koristi znatno nadmašuju štetu. **B je preporuka koju je bolje primenjivati nego ne primenjivati u praksi**, jer postoje najmanje pouzdani dokazi da intervencija poboljšava zdravstvene ishode i zaključuje se da koristi ipak nadmašuju štetu. **C znači da se ne može dati preporuka ni za, ni protiv rutinskog pružanja intervencije**. Postoje dokazi da intervencija može poboljšati zdravstvene ishode, ali se zaključuje da je ravnoteža koristi i štete preblizu da bi se dala bilo kakva preporuka. **D znači da se intervencija**

ne preporučuje rutinski u kliničkoj praksi. Pronađeni su ubedljivi dokazi da je intervencija neefikasna ili da šteta nadmašuje koristi. Najzad, ako se uopšte ne može formulisati preporuka jer su dokazi dvosmisleni ili međusobno suprotstavljeni, postavlja se oznaka **I** (od engleske reči „indeterminate“, tj. neodređeno).

Kada se izrađuju vodiči dobre kliničke prakse može se dogoditi da postoje propusti u primenjenoj metodologiji, ili da članovi komisije koja piše vodič nametnu svoje subjektivne poglede kao preporuke jer imaju predrasude ili određeni finansijski interes. Zato je neophodno pre prihvatanja preporuka iz vodiča proveriti da li je vodič urađen metodološki ispravno ili ne, i da li su sve preporuke zaista zasnovane, tj. treba proceniti njegov kvalitet. Za procenu kvaliteta vodiča dobre prakse koristi se upitnik **AGREE II,** koji procenjuje metodološku ispravnost, postupak razvoja vodiča i adekvatnost izveštavanja u vodiču. Ovaj upitnik ima 23 pitanja svrstana u 6 domena kojima se ispituju: obim i svrha vodiča, uključenost svih zainteresovanih strana, rigoroznost pretrage literature i selekcije kvalitetnih studija, jasnoća prikaza preporuka, primenljivost preporuka i nezavisnost komisije koja je pisala vodič od farmaceutskih kompanija.

FARMAKOEKONOMSKE STUDIJE

Ekonomija je nauka o tome kako društvo odlučuje šta će proizvoditi, kako će to činiti i za koga. Zdravstvena ekonomija je ekonomija primenjena na zdravstvo. Ona tretira zdravstvene usluge kao i ostale usluge na tržištu, uz uvažavenje specifičnosti. Farmakoekonomija je nauka koja proučava kako ljudi i društvo odlučuju da upotrebe ograničene resurse (koji su mogli da se upotrebe za nešto drugo) za nabavku lekova i medicinskih sredstava, i kako ih raspodeljuju ljudima i društvenim grupama da bi se dobilo na kvantitetu i kvalitetu života.

Postoje četiri vrste farmakoekonomskih studija:

1. Analiza minimiziranja troškova (engleski: Cost Minimisation Analysis) – primenjuje se kada je ishod dve alternativne terapije ISTI, pa se samo upoređuje njihova cena.

2. Analiza odnosa troškova i kliničkog efekta (engleski: Cost/Effectiveness Analysis) – primenjuje se u situacijama kada se ishod dva alternativna tretmana meri ISTIM "prirodnim" JEDINICAMA (npr. procenat sprečenih slučajeva infarkta mozga), a troškovi se mere novcem.

3. Analiza odnosa troškova i «korisnosti» (engleski: Cost/Utility Analysis) - upoređuje se "korisnost" dva različita tretmana (često različitih bolesti), uzimajući u obzir njihove troškove izražene u novcu. "Korisnost" se izražava najčešće dobijenim godinama života koje su korigovane za kvalitet (Quality Adjusted Life Years).

4. Analiza troškovi/dobit (engleski: Cost/Benefit Analysis) – izračunava se pridružena dobit svakog od tretmana (izraženo u novcu) i upoređuje sa njihovim troškovima. Tako je moguće upoređivati potpuno različite tretmane ("vrednost za novac").

Rezultati farmakoekonomske analize mogu jako varirati u zavisnosti od tačke gledišta. Npr. fondu zdravstvenog osiguranja neće odgovarati upotreba vrlo skupog leka koji će potpuno oporaviti pacijenta, ali društvu kao celini hoće, jer se smanjuju troškovi njegove nege i izdržavanja. Postoje tri osnovne tačke gledišta: institucionalna, društvena i individualna. Validna farmakoekonomska analiza obavezno definiše tačku gledišta sa koje se posmatraju troškovi i efekti.

Farmakoekonomska analiza kojoj se može verovati uzima u obzir sve raspoložive terapijske ili dijagnostičke alternative, kao i sve troškove koji nastaju prilikom lečenja ili dijagnostike. Postoje tri vrste troškova: **direktni** (troškovi samog leka, medicinske usluge, bolnički dan i sl.), **indirektni** (troškovi koje plaća pacijent ili društvo, npr. gubitak radnih dana, gubitak produktivnosti, i dr.) i **neuhvatljivi** (bol, strah, zabrinutost pacijenta). Validna farmakoekonomska analiza definiše koji troškovi su uzeti u obzir, i prema tome izvodi zaključak. Neuhvatljivi troškovi se obično ne uzimaju u obzir, jer ih je jako teško kvantifikovati.

O kliničkoj značajnosti rezultata farmakoekonomske studije se može govoriti ako su razlike u odnosu troškova i efekata između različitih terapijskih alternativa impresivne. Takođe, važno je da zaključak studije može da izdrži test promene polaznih parametara (npr. promene cene lekova, promene efikasnosti tretmana, i sl.). To utvrđujemo posebnom analizom rezultata studije, koju nazivamo «analiza osetljivosti».

Analiza osetljivosti je proces kojim se proverava tačnost i stabilnost rezultata i zaključaka farmakoekonomske analize, tako što se variraju vrednosti varijabli i polaznih pretpostavki u rasponu koji izgleda verovatan. Radi se zapravo o matematičkoj manipulaciji vrednostima polaznih varijabli i sticanju uvida u eventualne promene rezultata farmakoekonomske analize koje ove manipulacije izazivaju. Varijable koje se obično variraju su:

- Efektivnost/efikasnost leka
- Učestalost neželjenih dejstava
- Troškovi korišćenih resursa

Raspon u kome se menja vrednost varijabli se određuje tako što se utvrde gornje i donje granice, i to:

- Na osnovu prethodnog kliničkog istraživanja
- Na osnovu pregleda literature
- Na osnovu mišljenja onih koji donose odluke i eksperata zdravstvene službe

Rezultati farmakoekonomske studije se zatim ponovo izračunavaju koristeći postavljeni opseg vrednosti varijable. Može se pokazati da je jedna alternativa povoljnija u jednim uslovima, a druga alternativa u drugim. Treba odrediti tačku u kojoj jedan tretman postaje povoljniji od drugog, i proceniti verovatnoću pojave te tzv. prelomne tačke.

Analiza osetljivosti može biti **jednofaktorska**, gde se svaka varijabla koja se ispituje menja pojedinačno, i prati uticaj tih promena na ishode farmakoekonomske studije. Rezultat jednofaktorske analize osetljivosti se obično prikazuje pomoću tzv. “Tornado” dijagrama, koji je zapravo niz prečki postavljenih jedna ispod druge, a svaka prečka predstavlja promenu

ishoda farmakoekonomske studija kada se samo jedna određena ulazna varijabla menja**. Dvofaktorska ili višefaktorska** analiza senzitivnosti ispituju promene ishoda farmakoekonomske studije kada se dve ili više ulaznih varijabli menjaju istovremeno. Postoji i **probabilistička** analiza osetljivosti, kod koje se sve ulazne varijable menjaju istovremeno, i to po principu slučajnog unosa vrednosti iz prirodne distribucije svake od varijabli.

Međunarodni vodič za izveštavanje o sprovedenoj farmakoekonmskoj studiji se naziva CHEERS (Consolidated Health Economic Evaluation Reporting Standards), i može se naći na veb stranici https://www.equator-network.org/reporting-guidelines/cheers/ .

VRSTE KLINIČKIH STUDIJA

Mada je pethodno već bilo reči o randomiziranim, kontrolisanim kliničkim studijama, druge vrste kliničkih ispitivanja nisu bile pomenute. **Kliničke studije** nose takav naziv jer se sprovode na humanoj populaciji, tj. njihovi ispitanici su ljudi (pacijenti i/ili zdrave osobe). Ako se u studiji sprovodi neka intervencija, tj. ispituje se uticaj nekog leka, terapijskog ili dijagnostičkog postupka na zavisnu varijablu, i to se čini samo radi studije, takva klinička studija se naziva **INTERVENTNOM**. Interventne kliničke studije moraju biti kontrolisane (tj. intervencija mora imati nekog komparatora, tj. poredbenu grupu), randomizirane (tj. ispitanici se svrstavaju u ispitivanu i poredbenu grupu po principu slučajnosti) i dvostruko slepe (tj. ni istraživači, ni ispitanici ne smeju znati koji ispitanik pripada interventnoj, a koji kontrolnoj grupi), kako bi se izbegla pristrasnost. Druga vrsta kliničkih studija su **OPSERVACIONE**, kod kojih se ne primenjuje intervencija, već se samo posmatra i prati ono što se dešava sa zavisnom varijablom (ishodom) kod ispitanika koji su izloženi dejstvu neke nezavisne varijable (uzroka). Dok su kod kontrolisanih, randomiziranih kliničkih studija ispitanici veoma probrani (usled primene preciznih kriterijuma za uključivanje i isključivanje), pa je direktna primena njihovih rezultata na opštu populaciju upitna, kod opservacionih studija ispitanici su daleko

heterogeniji, više odražavaju opštu populaciju, pa se često opservacione studije nazivaju “studijama iz stvarnog života”. Postoje tri osnovne podvrste opservacionih kliničkih studija:

- Kohortne studije: ovde se ispitivana i kontrolna grupa formiraju prema izloženosti nezavisnoj varijabli, a zatim se pacijenti iz obe grupe prate tokom vremena i beleže ishodi koji kod njih nastaju. Ključna karakteristika kohortnih studija je dakle praćenje pacijenata tokom vremena, što znači da se pacijenti moraju ispitati bar dva puta, na početku i na kraju studije. Kohortne studije su dobile naziv po jedinici rimske vojske koja se zvala kohorta, i brojala je 600 vojnika; od početka služenja u vojsci članovi kohorte nisu napuštali svoju jedinicu, osim ako bi poginuli u borbi. Kohortne studije omogućavaju utvrđivanje incidencije ishoda.
- Studije preseka: i kod ove vrste opservacionih studija se ispitivana i kontrolna grupa formiraju prema izloženosti nezavisnoj varijabli, ali za razliku od kohortnih studija, u ovoj vrsti studija nema praćenja pacijenata u vremenu, već se posmatranje (opservacija) sprovodi samo jednokratno. Studije preseka ne mogu utvrditi incidenciju ishoda, ali zato mogu dati procenu njegove prevalencije.
- Studije slučaj/kontrola: u ovim studijama poredbene grupe se formiraju prema ishodu, slučajevi su ispitanici kod kojih se dogodio određeni ishod, a kontrole su ispitanici kod kojih se nije dogodio. Razlika između poredbenih grupa se potom traži u pogledu zastupljenosti nezavisne varijable, tj. uzroka.

Sve tri vrste opservacionih studija mogu prema načinu prikupljanja podataka biti retrospektivne ili prospektivne. Kod **retrospektivnih** studija podaci se prikupljaju iz medicinske dokumentacije, jer su se i uzroci i ishodi već dogodili u prošlosti. Kvalitet retrospektivnih studija direktno zavisi od kvaliteta i dostupnosti medicinske dokumentacije. Kod **prospektivnih** studija podaci o pacijentima se prikupljaju onda kada nastaju; prospektivne studije zahtevaju ulaganje mnogo više rada, ali su zato izbor varijabli koje se prate i tačnost prikupljenih podataka značajno veći.

DA LI JE U OBJAVLJENOJ STUDIJI KORIŠĆEN ADEKVATAN STATISTIČKI TEST ZA OBRADU REZULTATA?

Da bi rezultati studije bili metodološki ispravni (validni), neophodno je da se za njihovu obradu koriste odgovarajući statistički testovi. Važno je da dobro poznajemo korake koje treba učiniti prilikom izbora testa, kako ne bi napravili grešku i izabrali test koji se ne može upotrebiti za obradu podataka dobijenih studijom. Koraci u izboru statističkog testa se moraju praviti po sledećem redosledu:

1. Utvrditi šta je nezavisna (uzrok), a šta zavisna varijabla (ishod)
2. Definisati da li su varijable kontinualne (numeričke) ili kategorijske
    1. Kontinualna (numerička) varijabla (može uzeti bilo koju numeričku vrednost u određenom opsegu)
    2. Kategorijska varijabla (ako ima numeričkih vrednosti varijable, one mogu biti samo određene, tako da se jedinice posmatranja svrstavaju u kategorije koje se međusobno NE PREKLAPAJU)
        a. Kategorijske varijable imaju 2 nivoa merenja:
            i. Nominalno (npr. muški/ženski pol)
            ii. Ordinalno (kategorije su poređene po nekom redosledu, npr. osnovno, srednje i visoko obrazovanje)
        b. Kontinualne varijable takođe imaju dva nivoa merenja:
            i. Intervalno (jednaki intervali na nekoj skali, samo u određenom opsegu, ne postoji nula ni teoretski, npr. telesna temperatura)
            ii. U vidu odnosa (engl. „ratio“) – jednaki intervali na nekoj skali, teoretski može postojati nula a odnos između vrednosti može imati smisla, npr. serumski nivo CRP-a
    3. Izabrati između parametarskih i neparametarskih statističkih testova
        a. Parametarski testovi su mnogo senzitivniji, ali da bi se mogli koristiti, treba da budu ispunjene 4 pretpostavke:

i. Podaci moraju biti kontinualni (numerički)
ii. Jedinice posmatranja moraju biti izabrane slučajno iz populacije kojoj pripadaju (ponekada se može izostaviti)
iii. Podaci moraju biti normalno distribuirani (proverava se Kolmogorov–Smirnov ili Šapiro-Wilks testovima, koji ispituje da li se aktuelna distribucija značajno razlikuje od normalne)
iv. Varijabilnost podataka u grupama koje se porede treba da bude slična, tj. treba da postoji homogenost varijansi ili homoscedastičnost (proverava se Levene-ovim testom, koji ispituje da li su razlike u varijansama posledica slučajnosti, što će biti ako vrednost Levene-ovog testa odgovara verovatnoći nulte hipoteze većoj od 0.05)

4. Odrediti prema kojoj varijabli se formiraju grupe u studiji
    a. Kod eksperimentalnih, kohortnih i studija preseka nivoi nezavisne varijable određuju grupe
    b. Kod studija slučaj/kontrola nivoi zavisne varijable određuju grupe
5. Odrediti da li grupe sadrže različite (nezavisne) jedinice posmatranja tj. pacijente, pa se smatraju nezavisnim grupama ili se merenja vrše na istim jedinicama posmatranja u više navrata, pa su grupe zavisne
6. Koliko grupa se formira na osnovu mogućih vrednosti varijable: dve, više ili nije moguće svrstati jedinice posmatranja u grupe?

Kada se prođu svi navedeni koraci i nedvosmisleno odgovori na prethodna pitanja, test biramo prema principima prikazanim na sledećim tabelama:

Tabela 3: Testiranje razlike između dve grupe, kontinuirana ili ordinarna varijabla.

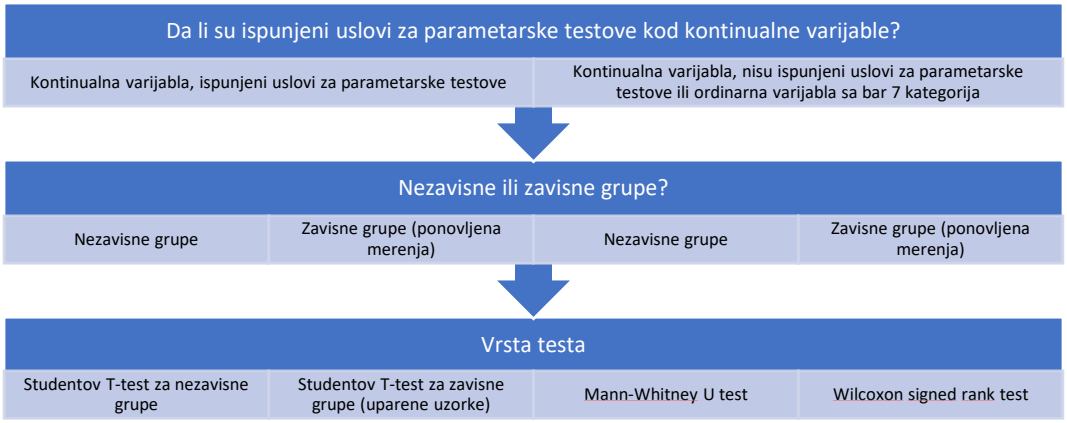


Tabela 4: Testiranje razlike između tri ili više grupa, kontinuirana ili ordinarna varijabla.

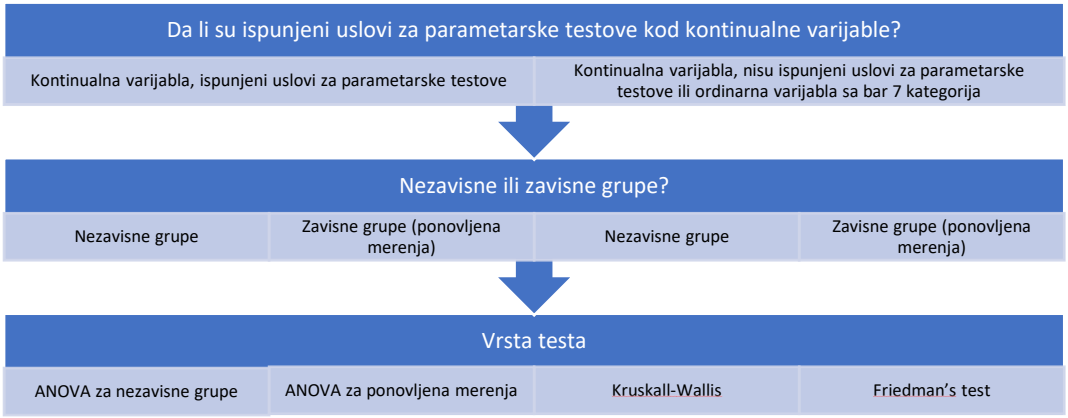


Tabela 5: Testiranje razlika između dve ili više grupa, kategorijska varijabla ili ordinarna varijabla sa manje od 7 kategorija.

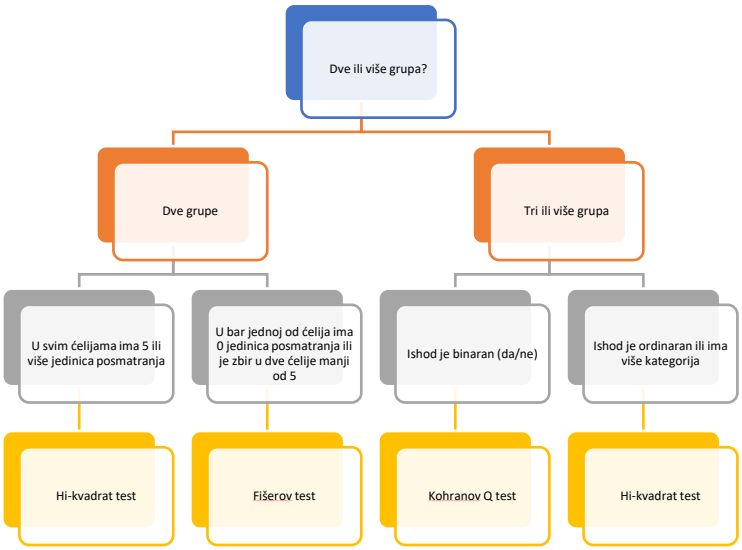


Tabela 6: Izbor testa kada postoje DVE NEZAVISNE varijable, a zavisna je normalno distribuirana.

Obe nezavisne varijable omogućavaju nezavisno merenje zavisne varijable	Jedna nezavisna varijabla omogućava ponovljeno merenje zavisne varijable, a druga nezavisno merenje	Obe nezavisne varijable omogućavaju ponovljeno merenje zavisne varijable
•Dvofaktorska analiza varijanse sa nezavisnim merenjima	•Dvofaktorska analiza varijanse sa mešovitim merenjima	•Dvofaktorska analiza varijanse sa ponovljenim merenjima

Tabela 7: Testovi za ispitivanje veze između zavisne i nezavisne varijable.

NAJČEŠĆE GREŠKE U STATISTIČKOJ ANALIZI REZULTATA

Greške u statističkoj obradi rezultata naučnih istraživanja imaju ozbiljne posledice: nastaju pogrešni zaključci koji dovode do neadekvatne upotrebe lekova, netačni rezultati koji precenjuju ili potcenjuju efikasnost odnosno bezbednost nekog terapijskog postupka, dolazi do navođenja drugih istraživača na pogrešan dizajn svojih studija i značajnog gubitka ljudskog rada i novca. Često već greške u dizajnu studije navode na pravljenje i statističkih grešaka, na primer nije jasno definisan ishod studije, veličina studijskog uzorka je nedovoljna, previše pacijenata je napustilo studiju pre njenog kraja, kod kliničkih studija nisu sprovedeni randomizacija i dvostruko zaslepljivanje, kontrolna grupa nije adekvatna, ili eksperimentalna i kontrolna grupa se razlikuju po većem broju parametara od kojih neki mogu uticati na ishod. Ipak, najveće posledice imaju greške u primeni statističkih

testova, tj. u samoj analizi podataka. Od grešaka u primeni statističkih testova, najvažnije je izbeći sledeće:

- Upotreba parametarskih testova kada ona nije moguća, tj. kada podaci nisu normalno distribuirani
- Korišćenje testa za nezavisne grupe kada su one zavisne, i obrnuto
- Ne koriste se multivarijantne metode kako bi se korigovali rezultati za uticaj zbunjujućih varijabli
- Ne primenjuje se korekcija za mnogostruka poređenja
- Neodgovarajuća post-hoc analiza podgrupa
- Nisu zadovoljeni uslovi za sprovođenje T-testa
- Nejednaka veličina uzoraka kod uparenog T-testa
- Nije primenjena Yates-ova korekcija kontinuiteta kada su grupe male
- Upotreba Hi-kvadrat testa kada je očekivani broj u ćeliji <5
- Podela kontinualnih podataka u ordinarne kategorije bez objašnjenja zašto i kako
- Upotreba linearne regresije, a da prethodno nije provereno da li je veza između nezavisne i zavisne promenljive linearna (pravi se grafik ostataka stvarnih i procenjenih vrednosti zavisne varijable iz same regresije – dobija se ravna linija ako postoji linearnost
- Upotreba binarne logističke ili Koks-ove regresije a da nije ispitano da li su zadovoljene pretpostavke za njihovu upotrebu (proporcionalnost hazarda za Koks-ovu regresiju i odsustvo multikolinearnosti i linearna veza između eksplanatornih varijabli i logit-a zavisne varijable kod logističke regresije)

Sledeća grupa grešaka u statistici se javlja prilikom opisa i/ili prikaza rezultata statističke obrade. Najpre se često ne navode svi testovi koji su korišćeni, da li su jednostrani ili dvostrani, da li u upareni ili ne. Ponekada autori navedu da su koristili neki statistički test koji nije prethodno validiran, i za koji nemaju objašnjenje. U prikazu deskriptivne statistike greška je

navesti srednju vrednost, ali ne i varijabilnost; srednja vrednost i standardna devijacija se mogu koristiti za opis samo podataka koji su normalno distribuirani – u protivnom se moraju koristiti medijana i interkvartilni opseg. Takođe, standardna greška se ne može koristiti umesto standardne devijacije za opis rezultata.

Prilikom prikaza rezultata statističke obrade neophodno je navoditi i intervale poverenja. Ako se prikazuje razlika između studijskih grupa, onda treba navesti interval poverenja za razliku između grupa, a ne za vrednosti upoređivane varijable u svakoj od grupa. Navođenje verovatnoće nulte hipoteze na sledeći način: “p = nije značajno”, “p <0.05” ili manje od neke druge arbitrarne granice, nije dozvoljivo – neophodno je navesti tačne p vrednosti, sa tri decimalne vrednosti (npr. 0,002). Prikazivanje numeričkih podataka sa besmislenom preciznošću je takođe greška (npr. srednji krvi pritisak jednako 99.347 mmHg).

Najzad, greške se prave i prilikom tumačenja rezultata. Ovde su navedena samo neka od pogrešnih tumačenja:

- Rezultat nekog statističkog testa “Nije značajan” se interpretira kao “nema efekta” ili “nema razlike”, a ustvari se ti ništa ne može zaključiti, jer rezultat može biti da nije značajan zbog malog uzorka.
- Ne razlikuju se pojmovi „statistički značajna razlika“ i „klinički značajna razlika“: rezultat može biti statistički značajan, a da nema nikakav klinički značaj. Na primer, statistički značajno sniženje holesterola u serumu sa 5.5 na 5.1 mM/L nema nikakav klinički značaj-
- Izvođenje zaključaka koji ne slede iz podataka dobijenih u studiji

METODOLOGIJA KRITIČKE PROCENE OBJAVLJENIH KLINIČKIH STUDIJA

Kada treba da procenimo metodološku ispravnost objavljene kliničke studije, korisno je da sledimo određeni algoritam, sa definisanim fazama.

Algoritam procene metodološke ispravnosti ima šest ključnih elemenata, tj. treba proceniti sledeće:

1. Interna validnost – da li je studija metodološki dobro urađena tako da nema pristrasnosti?
2. Interna konzistentnost – da li većina rezultata studije navodi na isti zaključak?
3. Eksterna validnost – stepen u kojem se rezultati studije mogu generalizovati na populaciju, posebno na populaciju iz koje je uzorak uzet.
4. Eksterna konzistentnost – da li se rezultati studije slažu sa rezultatima drugih istraživanja istog fenomena?
5. Klinički značaj
6. Preciznost dobijenih rezultata

Interna validnost se procenjuje na osnovu postojanja pristrasnosti (engl. bias-a) prilikom izrade kliničke studije, i to u nekoliko glavnih aspekata. Da bi studija imala internu validnost, neophodno je najpre da je **pristrasnost selekcije** minimalna ili odsutna, tj. test i kontrolna grupa treba da u svemu budu slične izuzev u tretmanima koji se porede (slične starosti, pola, trajanja, težine bolesti, itd.). Na taj način se umanjuje uticaj drugih (tzv. zbunjujućih) varijabli na vezu između tretmana koji se porede i ishoda lečenja. Druga pristrasnost koju treba eliminisati prilikom planiranja i sprovođenja kliničke studije je **pristrasnost u tretmanu**, tj. tretmani eksperimentalne i kontrolne grupe treba da se razliku samo po intervenciji koja se ispituje, a ne i po drugim elementima, kao što su trajanje izloženosti tretmanu, početak tretmana, primena dodatnih mera lečenja, itd. Ako se ovi drugi elementi razlikuju, nećemo biti sigurni na kraju studije da li je opservirani efekat rezultat same intervencije ili drugih elemenata. **Pristrasnost detekcije odnosno merenja efekta**: ako postoje razlike u detekciji ili merenju ishoda studije, npr. zbog neadekvatnog zaslepljivanja, zbog različite obuke istraživača ili drugih faktora, dobijene razlike u ishodima između grupa mogu biti posledica tih razlika, a ne efikasnosti samih intervencija. Najzad, treba izbeći i **pristrasnost osipanja**, koja se

odnosi na razliku između test i kontrolne grupe u broju pacijenata što napuštanju studiju pre njenog završetka.

Elementi interne (ne)validnosti se mogu prepoznati analizom poglavlja "Uvod" i "Metod" rukopisa koji opisuje sprovedenu studiju. Iz Uvoda publikovane kliničke studije treba izvući cilj studije, odnosno istraživačko pitanje (obično se navodi u poslednjoj rečenici Uvoda), a zatim proceniti da li ima sva četiri neophodna elementa (PICO – populacija, intervencija, kontrola i ishod). Takođe, da bi uvideli da li je cilj studije dobro postavljen, neophodno je da u Uvodu autori objasne šta se do sada zna o temi istraživanja, a šta je ostalo nejasno. Interna validnost studije zavisi, osim od pravilno postavljenog cilja, i od pravilnih postavki u odeljku "Metod". Prvo je neophodno da studijska populacija bude jasno definisana kroz kriterijume za uključivanje i isključivanje, a zatim da se iz nje izvuče adekvatan uzorak (ako je ikako moguće postupkom slučajnog izbora pacijenata za studiju) koji će biti dovoljno veliki da snaga studije (engl. "power") bude iznad 80%. Snaga studije je inače verovatnoća da će se studijom otkriti razlika između ispitivanih grupa koja zaista i postoji u stvarnosti. Sa snagom studije je komplementarna (sabira se do 100%) verovatnoća statističke greške drugog tipa, tj. verovatnoća da će se pokazati statističkim testom da između ispitivanih grupa nema razlike, a da ona zaista postoji. Dakle, ako snaga studije treba da bude veća od 80%, verovatnoća statističke greške drugog tipa treba da bude manja od 20%.

Veoma važni aspekti metodološke ispravnosti, tj. **interne konzistencije** kliničke studije jesu vrsta ishoda (zavisne varijable) i način njihovog merenja. Pravi ishodi kliničkih studija treba da budu preživljavanje, oporavak, izbegavanje hospitalizacije ili hirurške intervencije, ili slično. Pošto je za praćenje ovih ishoda potrebno dosta vremena (često nekoliko godina), istraživači pribegavaju praćenju surogat markera umesto pravih ishoda. Na primer, umesto da prate smrtnost kod lečenja dijabetesa, za šta su potrebne decenije, istraživači mere efekat antidijabetika na nivo glikoziliranog hemoglobina, što je surogat marker, koji nije u nedvosmislenoj vezi sa mortalitetom. Takođe je bitan način merenja ishoda studije: nije svejedno da li sam istraživač meri ishod ili to čini neko nezavisno od istraživača; važno je i da se merenje nekog labaratorijskog

parametra vrši validiranom metodom na atestiranom aparatu, a ne metodom ili na aparatu koji nikada nisu provereni ili kalibrisani.

Ukoliko studija ima više ishoda, onda se sprovodi i više statističkih testova, pa se povećava rizik od pronalaženja lažno pozitivnog rezultata (greška tipa I). Da bi se ovo uzelo u obzir i izbeglo, obično se primenjuju prilagođavanja poput Bonferonijeve korekcije ili kontrole stope lažnog otkrivanja. Bonferonijeva korekcija podrazumeva podešavanje kritičnog nivoa značajnosti od 0,05 deljenjem sa brojem izvršenih statističkih testova. Na primer, ako studija upoređuje promene u više parametara izvedenih iz srčane magnetne rezonance (npr. veličina infarkta, naprezanje miokarda i zapremine komora), prilagođavanja za višestruka poređenja mogu pomoći u izbegavanju lažno pozitivnih nalaza. Recimo da se poređenja vrše za čak 6 ishoda (parametara) iste studije, onda kritični nivo značajnosti verovatnoće nulte hipoteze uz Bonferonijevu korekciju bi bio 0.05/6 = 0.008. Dakle, samo ukoliko bi verovatnoća nulte hipoteze za bilo koje od ovih 6 poređenja bila manja od 0.008 moglo bi se reći da to poređenje pokazuje značajnu razlilku između studijskih grupa.

Drugi metod korekcije graničnih p vrednosti je Benjamini-Hochberg procedura za određivanje stope lažno pozitivnih rezultata (engl. „false discovery rate – FDR“) . Po ovoj proceduri najpre se izvršite svi planiran statistički testovi i zabeleže p-vrednosti za svako poređenje. Onda se poređaju sve p-vrednosti od najmanje do najveće. Najmanjoj p-vrednosti dodeljuje se rang jedan (i=1), sledećoj dva (i=2), i tako dalje, do poslednje koja ima rang m (i=m) (gde je (m) ukupan broj testova). Za svaku rangiranu p-vrednost izračunava se odgovarajuća Benjamini-Hochberg kritična vrednost (prag značajnosti) korišćenjem formule: $Kritična\ vrednost = \frac{i}{m} * Q$, gde je:(i) - rang p-vrednosti (od 1 do (m)) (m) - ukupan broj testova (hipoteza). a (Q) - željena stopa lažnih otkrića (obično 0.05; 0.01 ako hoćemo da budemo stroži; 0.1 ako se radi o pilot studiji). Zatim se pronalazi najveći rang (k) za koji je originalna p-vrednost (P_k) manja ili jednaka izračunatoj kritičnoj vrednosti Odbacuju se (proglašavaju značajnim) sve nulte hipoteze čije su p-vrednosti u rangu od 1 do (k) (uključujući i (k)).

Eksterna validnost zavisi od reprezentativnosti uzorka za populaciju iz koje je uzet, tj. od stepena u kome se rezultati dobijeni na uzorku mogu

generalizovati na celu populaciju. Postoji više faktora koji mogu da utiču na reprezentativnost uzorka. Najpre sam process pristanka za učešće u studiji ima značajan uticaj: pacijenti koji pristaju da uzmu učešća u studiji mogu biti samo oni sa određenim stadijumom bolesti i sa određenim psihološkim karakteristikama, dok drugi "tipovi" pacijenata mogu ostati izvan studije. Kriterijumi za uključivanje i isključivanje pacijenata u studiju mogu biti ponekada postavljeni tako da u okviru ciljne populacije vrše selekciju pacijenata, pa uzorak neće biti reprezentativan. Stavovi pacijenata o lečenju uopšte takođe mnogo utiču na izbor ispitanika. Na primer, ukoliko je preovlađujući stav pacijenata da hirurške intervencije treba izbegavati, neka klinička studija iz oblasti hirurgije će uključiti samo pacijente koji ne zaziru od hirurške intervencije, tj. uzorak za studiju neće biti reprezentativan za celu ciljnu populaciju. Manje važni faktori koji mogu mnogo uticati na reprezentativnost uzorka, tj. na eksternu validnost studije su: razlike u zdravstvenim sistemima zemalja u kojima se sprovodi multicentrična studija, razlike u stepenu ekspertize istraživača, korišćenje surogat markera (npr. laboratorijskih rezultata) kao ishoda studije umesto pravih ishoda (izlečenje ili mortalitet).

Bornhöft i saradnici su formirali listu pitanja kojima se ispituje mera eksterne validnosti publikovane kliničke studije, od kojih se ovde navode samo najvažnija: (1) u kojoj meri kriterijumi za uključivanje i isključivanje definišu populaciju; (2) da se dijagnostička procedura primenjena u studiji inače koristi u praksi; (3) da li uzorak odgovara populaciji po starosti, polu, težini bolesti, konkomitantnim oboljenjima, konkomitantnoj terapiji, itd; (4) da li se tretman primenjen u studiji koristi i u rutinskoj kliničkoj praksi na isti način, od strane osoba sličnih kvalifikacija; (5) da li su izabrani ishodi relevantni za praksu i pacijenta; (6) da li su sledeći važni ishodi uzeti u obzir: kvalitet života, subjektivno zdravlje, opšte procene pacijenta, razlozi za prekid lečenja, upotreba pratećih tretmana, povratni efekat nakon prekida lečenja; (7) da li je prikupljeno dovoljno podataka da se intraindividualna varijabilnost ispita; (8) da li stopa napuštanja studije odražava rutinsku praksu.

Eksterna konzistencija rezultata studije se utvrđuje poređenjem njenih rezultata sa rezultatima drugih studija. Da bi se to učinilo na sistematičan način, i da se ne bi izostavile neke studije, neophodno je da se

eksterna konzistencija utvrdi putem izrade sistematskog preglednog članka. Sistematski pregled literature sa istom temom kao što je tema studije čiju eksternu konzistenciju želimo da utvrdimo treba da obezbedi da se: (1) obuhvate sve publikovane studije; (2) u obzir uzmu svi ishodi studije (i primarni, i sekundarni); (3) ispita metodološka ispravnost, tj. kvalitet svih obuhvaćenih studija; (4) proceni stepen eksterne konzistencije, ili (što je druga strana medalje) stepen heterogenosti rezultata. Ako rezultati ne odstupaju mnogo, heterogenost će biti mala (meri se Kohranovim Q-testom i I^2 analizom). Kohranov Q test se sastoji u izračunavanja veličine odstupanja rezultata svake pojedinačne studije obuhvaćene sistematskim pregledom od prosečnog rezultata svih studija, pri čemu se svako odstupanje ponderiše varijabilnošću zabeleženom u odgovarajućoj studiji. Izračunata Q vrednost se distribuira po Hi-kvadrat distribuciji, tako da njene ekstremne vrednosti (koje spadaju u vrednosti ukupno manje učestalosti od 10%) znače da je heterogenost isuviše velika, tj. da nema eksterne konzistencije. Q vrednost se izračunava na sledeći način: $Q = \sum w_i * (y_i - \bar{y})$, pri čemu je w_i recipročna vrednost varijanse ($\frac{1}{v_i}$), y_i – veličina efekta studije i, $\bar{y}$ - prosečni efekat svih studija.

I^2 analiza podrazumeva da se najpre izračuna I^2 vrednost, koja odražava uticaj broja studija u okviru sistematskog pregleda na Q vrednost, a zatim rezultat protumači. Ako je Q vrednost jednaka ili manja od broja studija umanjenog za 1, onda je I^2 vrednost jednaka nuli. Ako je pak Q vrednost veća od broja studija umanjenog za 1, onda se I^2 vrednost izračunava na sledeći način: $I^2 = \frac{Q-(k-1)}{Q} * 100\%$. Ako je I^2 vrednost manja od 40%, onda je heterogenost mala, tj. postoji eksterna konzistencija rezultata; visoka heterogenost je ako je I^2 vrednost veća od 75%.

Klinički značaj rezultata kliničke studije se procenjuje u zavisnosti od tipa ishoda. Ako je ishod kategorijska veličina, npr. pacijent preživeo ili nije preživeo, klinički značaj se procenjuje na ranije pomenuti način – izračunavanjem smanjenja apsolutnog rizika i broja bolesnika koji treba lečiti novom terapijom da bi jedan pacijent više preživeo. Ako je međutim ishod studije kontinualna varijabla, npr. dijastolni krvni pritisak, klinički značaj se utvrđuje pomoću koncepta nazvanog “minimalna klinički važna

razlika". Minimalna klinički važna razlika je promena u vrednosti kontinualnog ishoda koja dovodi do smanjenja rizika od ozbiljnih loših posledica po pacijenta, kakvi su smrtni ishod, hospitalizacija, hirurška intervencija, invaliditet ili pojava malignih oboljenja. Na primer, za smanjenje dijastolnog krvnog pritiska od 10 mmHg je poznato da smanjuje smrtnost pacijenata u desetogodišnjem periodu, što ga čini klinički značajnim. Pošto takav efekat nije pokazan za smanjenje dijastolnog pritiska od 9 mmHg, to znači da je smanjenje dijastolnog krvnog pritiska od 10 mmHg **minimalna** klinički važna razlika. Dakle, ako neki lek pokaže da snižava dijastolni krvni pritisak za više od 10 mmHg, pri čemu je i donja 95%-tna granica poverenja iznad 10 mmHg, može se reći da taj lek ima klinički značajan efekat.

Preciznost dobijenih rezultata kliničke studije se meri veličinom intervala poverenja: što je 95%-tni interval poverenja uži, rezultati iz studijskog uzorka predstavljaju precizniju procenu vrednosti odgovarajuće varijable u populaciji iz koje je uzorak uzet ili razlike u toj varijabli između grupa koje se porede u studiji. Interval poverenja se sužava (tj. rezultati postaju precizniji) kada se povećava veličina studijskog uzorka, i obrnuto. Iz intervala poverenja koji je objavljen u studiji moguće je izračunati kolika je snaga studije, tj. verovatnoća da će studija otkriti razliku koja zaista postoji između poredbenih grupa. Ako je snaga studije manja od 0.8 (ili 80%, što je isto ako verovatnoću izrazimo u procentima), rezultati studije nisu dovoljno precizni da bi im se moglo verovati.

SISTEMATSKA GREŠKA U SPROVOĐENJU KLINIČKIH STIDIJA ILI PRISTRASNOST (ENGL. "BIAS")

Prilikom obrade rezultata kliničkih studija moguće su dve vrste grešaka: **slučajna statistička greška**, koja nastaje usled nedovoljnog broja ispitanika ili prevelike varijabilnosti rezultata, i **sistematska greška**, koja nastaje zbog dejstva nekog faktora koji utiče na rezultate u nekom smeru, a

prisutan je zbog greške u dizajnu (postavci) same studije. Postoje tri vrste slučajne statističke greške:

- Greška tipa I : verovatnoća da će se posle obrade rezultata pokazati da POSTOJI značajna razlika među upoređivanim grupama, a da ona zaista ne postoji.
- Greška tipa II : verovatnoća da će se posle obrade rezultata pokazati da NE postoji značajna razlika među upoređivanim grupama, a da ona zaista postoji.
- Greška tipa III : nastaje kada se korektno odbaci nulta hipoteza, ali je sama hipoteza bila pogrešno postavljena, tj. njeno odbacivanje nam daje pogrešan zaključak o stvarnom stanju u populaciji.

Nastanak slučajnih statističkih grešaka se može izbeći pažljivim izračunavanjem potrebne veličine uzorka pacijenata za studiju, tako da verovatnoća njihove pojave bude dovoljno mala. Obično se za dovoljnu veličinu uzorka pacijenata smatra ona koja obezbeđuje da je verovatnoća nastanka greške tipa I manja od 0.05, a verovatnoća nastanka greške tipa II manja od 0.2. Često se umesto željene verovatnoće nastanka greške tipa II koristi njena komplementarna vrednost, tj. verovatnoća da će se između upoređivanih grupa naći razlika koja zaista POSTOJI. Ovu poslednju verovatnoću nazivamo **snaga studije** (engl. "power of study"), i ona mora biti veća od 0.8 da bi se izbegla greška tipa II.

Sistematske greške se mogu izbeći samo dobrim dizajnom studije. Neke od najčešćih sistematskih grešaka (drugi naziv je pristrasnost) su:

- **Selekciona pristrasnost**, sa više podtipova:
    - Podtip pristrasnosti „Incidencija-prevalencija" (sinonimi: Neyman-ova pristrasnost ili pristrasnost u pogledu preživljavanja). Ova pristrasnost se javlja kod studija preseka kada pokušamo da procenimo rizik od neke bolesti na osnovu podataka prikupljenih u jednoj vremenskoj tački, a ne na osnovu podataka prikupljenih tokom određenog vremenskog perioda u grupi izloženih pacijenata. Kod studija slučaj/kontrola može biti pristrasan izbor slučajeva za koji se zna da su bili izloženi faktoru rizika koji je predmet istraživanja.

- Još jedan podtip selekcione pristrasnosti je „pristrasnost zbog spektra bolesti“. Kod ovog podtipa, ukoliko se u studiju uključe samo pacijenti u ranoj ili samo pacijenti u kasnoj fazi bolesti, preživljavanje i drugi ishodi će biti sasvim drugačiji.
- Pristrasnost upućivanja se javlja kada je pojedinac ili navike lokalne prakse određuju koji subjekti se podvrgavaju određenoj studiji ili ako samo određena populacija pacijenata ima pristup mestu ispitivanja. Tako će postojati razlike između proučavane populacije i ciljne populacije.
- Pristrasnost učešća nastaje zbog faktora koji utiču na konačnu odluku pacijenata da li će prihvatiti učešće u studiji. Nikada neće svi pacijenti koji ispunjavaju kriterijume za uključivanje pristati da učestvuju, iz različitih razloga, od kojih neki mogu biti banalni: nedostatak prevoza, vreme vizita koje ne odgovara pacijentu, loša verbalna ili neverbalna kominikacija istraživača koji uzima informisani pristanak od pacijenta, itd.

- **Pristrasnost detekcije** (sinonim: informaciona pristrasnost):
  - Podtip „neodgovarajuća klasifikacija“. Pristrasnost pogrešne klasifikacije nastaje kada je proces otkrivanja statusa izloženosti (pristrasnost identifikacije izloženosti) i/ili ishoda (pristrasnost identifikacije ishoda) nesavršen, tj.izložene/obolele osobe su pogrešno klasifikovane kao neeksponirane/neobolele, i obrnuto. Prema tome kako se vrši identifikacija izloženosti ili ishoda, pristrasnost pogrešne klasifikacije može biti:
    - Pristrasnost zbog slabog sećanja ispitanika kod retrospektivnih studija (engl. „recall bias“)
    - Pristrasnost osobe koja intervjuiše ispitanike (engl. „interviewer bias“)
    - Pristrasnost osobe koja meri neki od ishoda (engl. „observer bias“) – da bi se sprečio, neophodno je da sva merenja u studiji sprovode osobe nezavisne od istraživača
  - Fenomen „regresije ka srednjoj vrednosti“. Poznato je da promenljiva koja pokazuje ekstremnu vrednost na svom prvom merenju će imati tendenciju da bude manje ekstremna

u narednim procenama. Ovaj fenomen je prisutan kod longitudinalnih studija koje istražuju povezanost između merenja kontinualne varijable i rizika za nastanak određenog ishoda.

    - „Pristrasnost vremena ulaska u studiju u odnosu na početak bolesti“ (engl. „lead time“ bias). Ovaj podtip pristrasnosti se javlja kod prospektivnih studija koje se bave procenom efikasnosti skrininga. Ko dove pristrasnosti na različite pacijente se primenuju različiti kriterijumi za dijagnozu bolesti (tj. dijagnoza može biti postavljena u latentnom periodu ili u ranoj fazi, naspram dijagnoze postavljene u poodmakloj fazi). Preživljavanje pacijenata dijagnostikovanih u ranoj fazi bolesti je znatno duže, pa možemo pogrešno zaključiti da je rana dijagnoza zbog skrininga korisna.

- **Pristrasnost osipanja**, ili „izgubljen za praćenje“ (engl. “attrition”), se javlja u prospektivnim kohortnim studijama kada pojedinci izgubljeni za dalje praćenje tokom studije nemaju istu verovatnoću kliničkog ishoda od interesa za poređenje sa pojedincima koji ostaju na posmatranju do kraja studije.
- **Pristrasnost zbog razlika u tretmanu pacijenata.** Ovaj vid pristrasnosti je veoma čest u kliničkim studijama, i odnosi se na pojavu da poredbene grupe nisu tretirane na isti način ne samo u pogledu primene leka ili medicinskog sredstva koje se ispituje, već i u pogledu ishrane, navika, upotrebe lekova koji se izdaju bez recepta ili dodataka ishrani, ili fizičke aktivnosti. Sve ove dodatne faktore nije lako kontrolisati u kliničkoj studiji, pa se često u publikacijama koje prikazuju rezultate kliničkih studija i ne pominju. Međutim, uticaj pomenutih faktora može biti veoma veliki, i ponekada nadmašiti uticaj leka koji se ispituje. Upravo iz tih razloga, pri analizi svake publikovane kliničke studije treba pažljivo ispitati da li postoji razlika u tretmanu poredbenih grupa osim primene intervencije koja se ispituje.
- **Pristrasnost u izveštavanju** se sastoji u tome da autori u publikaciji koja opisuje kliničku studiju izostave “negativne” rezultate, tj. ona poređenja gde se nije pokazala statistički značajna razlika između grupa. Takvim postupkom se prikriva istina i čitaoci publikacije

navode na pogrešan zaključak da je intervencija koja je ispitivana u studiji efikasna. Takođe, prilikom obrade rezultata studije moraju se uzeti u obzir svi pacijenti koji su ušli u studiju: tzv. **„intention-to-treat“ analiza.** Ovakva analiza se vrši radi prevencije kršenja prvobitne randomizacije i radi uzimanja u obzir svih pacijenata koji su napustili studiju ili prešli na alternativno lečenje. Ako se uzmu u obzir samo pacijenti koji su završili studiju, to se naziva **„per protocol“** analiza. Ovakva analiza predstavlja uvođenje pristrasnosti, tj. sistematske greške.

NEDOSTAJUĆI PODACI

Često u toku izvođenja studije istraživači ne prikupe sve potrebne podatke za svakog pacijenta, tj. podaci za neke varijable kod pojedinih pacijenata nedostaju. Na primer, u studiji koja ispituje efekat nekog antihipertenziva, od pacijenata izloženih tom leku treba prikupiti osim demografskih podataka i vrednosti krvnog pritiska i mnoge laboratorijske rezultate, kao što su nivo kreatinina, broj eritrocita, C-reaktivni protein, nivo albumina, itd. Ako nekom pacijentu kreatinin nije izmeren, taj podatak nedostaje, i onda se takav pacijent ne može uključiti u statističku obradu rezultata. Ako je takvih pacijenata puno, dolazi se u opasnost da će rezultati statističke obrade biti nesigurni, jer je malo pacijenata sa svim kompletnim podacima koji se mogu statistički obraditi. U takvoj situaciji, neki istraživači pribegavaju dopuni nedostajućoih podataka.

Dopunu nedostajućih podataka je opravdano učiniti ako su podaci ***potpuno nasumično nedostajući***, što zapravo znači da je verovatnoća nedostatka nekog podatka nezavisna od svih uočenih i neuočenih karakteristika pacijenata. Ako su podaci ***delimično nasumično nedostajući***, tj. ako postojanje ili nepostojanje nekog podatka zavisi od posmatranih varijabli, dopunu podataka je opravdano učiniti samo uz prilagođavanja korišćenjem statističkih modela (npr. višestruke imputacije), što donekle može ublažiti pristrasnost. Najzad, ako ***nedostajući podaci nisu nasumični***, što znači da su nedostajući podaci povezani sa neuočenim ishodima (npr.

pacijenti sa određenim stanjem možda radije ne bi odgovarali na pitanje upitnika u vezi sa tim stanjem), onda nije opravdano vršiti dopunu nedostajućih podataka.

Dopuna nedostajućiih podataka se može obaviti na nekoliko načina. Najjednostavniji način je **unošenje srednje vrednosti** studije za nedostajuće podatke. Ovakav metod međutim veštački smanjuje varijabilnost podataka i ignoriše povezanost ishoda sa određenim varijablama. **Metod uslovne imputacije vrednosti**, gde se regresioni model koristi za izračunavanje nedostajućih vrednosti na osnovu drugih varijabli, takođe nije idealan, jer pojačava veštački povezanost određenih varijabli i ishoda. **Višestruko unošenje nedostajućih podataka** je za sada najbolji metod. Prvo, za svaku promenljivu koja ima nedostajuće vrednosti, algoritam nasumično popuni prazna mesta srednjim vrednostima. Zatim, za svaku promenljivu posebno, algoritam izračunava nedostajuće vrednosti na osnovu multivarijantnog regresionog modela koji uključuje sve varijable. Na osnovu poređenja srednjih vrednosti i vrednosti dobijenih regresijom, algoritam dalje procenjuje koje vrednosti su realnije i zamenjuje prvobitno popunjene vrednosti boljim procenama. Ovaj proces se ponavlja više puta u ciklusima (5-20 puta) da bi se dobile što bolje procene nedostajučih podataka. Na kraju, ceo postupak se ponavlja nekoliko puta, kako bi se dobilo više različitih verzija popunjenih podataka, što omogućava naknadnu statističku analizu.

PROCENA VELIČINE EFEKTA U KLINIČKIM STUDIJAMA

Veličina efekta novog metoda lečenja u kliničkoj studiji se procenjuje na nekoliko načina, u zavisnosti od tipa ishoda. Ako je ishod kategorijska varijabla (npr. ozdravio/nije ozdravio ili umro/nije umro), efekat se meri odnosom šansi (engl. “odds ratio”) ili smanjenjem apsolutnog i relativnog rizika. S druge strane, ako je ishod studije kontinualna varijabla, veličina efekta se meri sa Koenovim D indeksom.

Apsolutno smanjenje rizika /ARR) od nepovoljnog ishoda je razlika učestalosti nepovoljnog ishoda u grupi kontrolnih pacijenata i grupi pacijenata sa ispitivanim lekom: ARR = R1 – R. Relativno smanjenje rizika je apsolutno smanjenje rizika podeljeno sa učestalošću nepovoljnog ishoda u grupi sa poredbenim lekom: RRR = ARR / R1. Broj bolesnika koje treba lečiti kako bi se izbegao jedan neovolan ishod se dobija kada se broj 100 podeli sa apsolutnim smanjenjem rizika: (NNT) = 100 / ARR. Odnos šansi ("odds ratio" – OR) je količnik šansi ("odds"-a) da dođe do određenog ishoda u grupi pacijenata sa ispitivanim lekom i grupi kontrolnih pacijenata. Na primer, zamislimo da je rezultat neke studije prikazan u sledećoj tabeli.

Tabela 8. Procena veličine efekta kada je ishod studije kategorijski.

	Umrli	**Preživeli**
Primali lek A	5	40
Nisu primali lek A	15	60
UKUPNO	20	100

ODDs za smrtni ishod u grupi na leku A je: 5/40 = 0.125 (to je šansa da će osoba koja prima lek A umreti)

ODDs za smrtni ishod u grupi bez leka A je: 15/60 = 0.25 (to je šansa da će osoba koja ne prima lek A umreti)

ODDs RATIO (OR) = 0.125 / 0.25 = 0.5

Interpretacija: osobe koje primaju lek A imaju upola manju šansu da umru nego osobe koje ne primaju lek A

Granice poverenja (engl. "confidence interval" – CI) su krajnje tačke opsega u kome se nalazi određeni procenat vrednosti ODDS RATIO-a u populaciji. Granice poverenja 95% su krajnje tačke opsega u kome se nalazi 95% vrednosti ODDS RATIO-a u populaciji. Formula za izračunavanje 95%-tnih granica poverenja je:

$$CI = e^{\left[\ln(OR) \pm 1.96 * \sqrt{\left(\frac{1}{5}+\frac{1}{15}+\frac{1}{40}+\frac{1}{60}\right)}\right]} = 0.17 - 1.48$$

Koenov D je standardizovana veličina efekta za merenje razlike između srednjih vrednosti kontinualne varijable dve grupe. Često se koristi kada se poredi tretman sa kontrolnom grupom, gde je ishod kontinualna varijabla, npr. krvni pritisak, broj leukocita, nivo kreatinina, itd. Adekvatan je za analizu rezultata Studentovog t-testa i jednofaktorske analize varijanse. Koenov D se izračunava na sledeći način:

$$Koenov\ D = \frac{(srednja\ vrednost\ u\ eksp.grupi - srednja\ vrednost\ u\ kontrolnoj\ grupi)}{zbirna\ standardna\ devijacija\ u\ obe\ grupe}$$

Ako je Koenov D oko 0.2, efekat je mali, a ako je oko 0.5, efekat je umeren. Za efekat se može reći da je veliki ako je Koenov D 0.8 ili više.

ISHODI KLINIČKIH STUDIJA

Kliničke studije obično imaju više ishoda, od kojih je jedan primarni (ili glavni) ishod, a ostali su sekundarni. I primarni, i sekundarni ishodi mogu biti kategorijski ili kontinualni. Kvalitetne kliničke studije kao glavni imaju neki od **ozbiljnih ishoda** (engl. "hard outcomes"), poput smrti ili stanja koja imaju trajne posledice po pacijenta (npr. infarkt miokarda, infarkt mozga, izlečenje pneumonije, itd.). Takav pristup je najbolji, jer bez prikrivanja ukazuje na efekte intervencije (leka ili drugog tretmana) u studiji. Međutim, istraživači ponekada u nemogućnosti da zbog nedovoljnog broja pacijenata pokažu efekte ispitivane intervencije na ozbiljne ishode pribagavaju korišćenju složenih (kompozitnih) ishoda ili surogat markera. **Složeni (kompozitni ishodi)** se sastoje od više ishoda, na primer složeni ishod koga čine smrt, infarkt miokarda ili infarkt mozga. Kada u studiji nema dovoljno pacijenata koji su imali samo jedan od ovih ishoda, studija nema dovoljnu statističku snagu da pokaže efekat intervencije; sa uvođenjem složenog ishoda broj pacijenata sa ishodom se povećava i studija dobija potrebnu statističku snagu da pokaže efekat intervencije. Loša strana složenih ishoda

je u tome da nije jasno koliki je doprinos svakog od ovih ishoda u pokazanom efektu intervencije, pa je tumačenje rezultata ograničeno.

Surogat markeri su obično laboratorijski podaci koji su indirektno povezani sa ozbiljnim ishodima lečenja, pa ukoliko se pokaže da ispitivana intervencija u studiji utiče na surogat marker, može se pretpostaviti da bi uticala i na ozbiljne ishode. Mada su surogat marker korisni, i podesni za pokazivanje efekata, takvi efekti ne mogu ništa reći o kliničkom značaju intervencije, jer efekat na ozbiljne ishode i dalje ostaje nepotvrđen.

U poslednje vreme se kod studija sa većim brojem ishoda koristi tzv. **hijerarhijska (ili višeslojna) analiza.** Ishodi se procenjuju po unapred definisanom redosledu, osiguravajući da se kritičniji ishodi (npr. mortalitet) analiziraju pre manje kritičnih (npr. ublažavanje simptoma). Ova metoda izbegava statističke probleme povezane sa višestrukim poređenjima i osigurava da se efekti lečenja procenjuju na strukturiran način. Jedna od metoda hijerarhijske analize je **"odnos pobeda i poraza".** Po ovoj metodi najpre se više ishoda rangira po kliničkom značaju. Parovi pacijenata (jedan iz interventne grupe i jedan iz kontrolne grupe) se porede sekvencijalno na osnovu hijerarhije ishoda. Ako pacijent u interventnoj grupi ima povoljniji ishod najveće kritičnosti, to se smatra „pobedom". Ako kontrolni pacijent ima bolji ishod, to je „poraz". Ako nema razlike, poređenje se nastavlja na sledećem rangiranom ishodu. Odnos pobeda se zatim izračunava kao ukupan broj pobeda podeljen sa ukupnim brojem poraza. Odnos pobeda veći od 1 sugeriše da je intervencija efikasnija od lečenja u kontrolnoj grupi. Odnos pobeda oko jedinice ukazuje na malu razliku između grupa.

Metoda "odnos pobeda i poraza" pruža veću osetljivost u otkrivanju koristi od lečenja uključivanjem kliničkog značaja svakog ishoda. Posebno je korisna u studijama hroničnih bolesti, gde su složeni ishodi uobičajeni, a mortalitet, hospitalizacije i ublažavanje simptoma treba hijerarhijski proceniti.

STUDIJE DOKAZIVANJA EKVIVALENTNOSTI ILI NE-INFERIORNOSTI NOVE TERAPIJE U ODNOSU NA STANDARDNU

Kada je nova terapija u odnosu na standardnu efikasnija, onda je potrebno u kliničkoj studiji to dokazati, pa se takve studije nazivaju "studije superiornosti". Ako pak nova terapija ne donosi veću efikasnost u odnosu na standardnu, ali ima manje neželjenih dejstava, dovoljno je u kliničkoj studiji pokazati da nema razlike u efikasnosti nove i standardne terapije ili bar da nova terapija nije manje efikasna od standardne. Takve kliničke studije se nazivaju "studije ekvivalentnosti", odnosno "studije ne-inferiornosti".

Klasična randomizirana klinička studija je zapravo ispitivanje superiornosti: da li je eksperimentalna/nova terapija superiornija od utvrđene referentne/standardne ili placebo terapije u poboljšanju ishoda lečenja. U okolnostima kada se superiornost nove terapije ne može dokazati, sledeća najbolja strategija bi bila da se proceni da li je nova terapija ekvivalentna ili u najmanju ruku „nije neprihvatljivo gora" od standardne terapije.

Cilj ispitivanja ekvivalencije je da se pokaže da je nova terapija ekvivalentna sa standardnom terapijom korišćenjem granice sa obe strane nulte razlike u efektu koja prikazuje maksimalnu podnošljivu razliku za ekvivalentnost. Ako izračunati interval poverenja razlike u ishodima nove i standardne terapije u kliničkoj studiji leži striktno unutar prethodno definisanih granica, kaže se da je nova terapija ekvivalentna sa starom.

S druge strane, u studijama ne-inferiornosti, nova terapija nije neprihvatljivo gora (tj. neinferiorna) od standardne ako donja granica izračunatih 95%-tnih granica poverenja oko razlike između efekata nove i standardne terapije ne prelazi unapred definisanu granicu (marginu) ne-inferiornosti.

Važan korak u dizajniranju kliničkih studija ne-inferiornosti je izbor granice ne-inferiornosti. Polazi se ili od jedne placebo-kontrolisane studije standardne terapije ili meta-analize da bi se dobila 95%-tna granica poverenja oko izračunate razlike u ishodu između standardne terapije i placeba. Donja granica 95%-tnih granica poverenja ili granica poverenja

najbliža nikakvom efektu se obično uzima za osnovu određivanja granice ne-inferiornosti. Samu granicu ne-inferiornosti izračunavamo kao 50% donje 95%-tne granice poverenja.

ALTERNATIVA MEDICINSKIM NAUKAMA ZASNOVANIM NA DOKAZIMA

Ako ne bi lečili i dijagnostikovali naše pacijente prema principima MZD, šta bi bila alternativa? Možemo rešavati samo probleme pacijenata sa kojima imamo iskustvo, što je uvek ograničeno, jer naše lično iskustvo ne može biti veliko. Sve zdravstvene probleme pacijenata sa kojima nemamo iskustvo nećemo moći da zbrinemo, i onda nam ostaje ili da pacijenta otpustimo bez rešenja (a njegova patnja se nastavlja), ili da ga pošaljemo drugom lekaru za koga pretpostavljamo da bi mogao rešiti taj problem. Dok je prva mogućnost moralno, stručno i administrativno neprihvatljiva, druga mogućnost će imati pozitivan ishod samo ako drugi lekar primenjuje principe MZD u svom radu! Dakle, ako bolje pogledamo, zapravo ne postoji racionalna alternativa upražnjavanju MZD u kliničkoj praksi.

LITERATURA

1. Sackett D. Evidence-based Medicine. NHS R&D Centre for Evidence-Based Medicine, Harcourt Brace, Oxford, U.K., 1997.
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2.
3. Bettany-Saltikov J, Whittaker VJ. Selecting the most appropriate inferential statistical test for your quantitative research study. Journal of Clinical Nursing. 2014;23(11-12):1520-31.
4. Lang T. Twenty statistical errors even you can find in biomedical research articles. Croat Med J. 2004;45(4):361-70.

5. Masic I, Jankovic SM, Begic E. PhD Students and the Most Frequent Mistakes During Data Interpretation by Statistical Analysis Software. Stud Health Technol Inform. 2019;262:105-109.
6. Jankovic SM, Masic I. Methodological Errors in Clinical Studies Published by Medical Journals of Ex-Yugoslav Countries. Acta Inform Med. 2020;28(2):84-93.
7. Masic I, Jakovljevic M, Sinanovic O, Gajovic S, Spiroski M, Jusufovic R, Sokolovic S, Prnjavorac B, Zerem E, Djulbegovic B, Porovic S, Jankovic S, Hadzikadic M, Zunic L, Begic E, Nislic E, Begic N, Becirovic E, Cerovac A, Skrijelj V, Nuhanovic J. The Second Mediterranean Seminar on Science Writing, Editing and Publishing (SWEP - 2018), Sarajevo, December 8th, 2018. Acta Inform Med. 2018;24(4):284-299.
8. Masic I, Jankovic SM. Comparative Analysis of Web of Science and Pubmed Indexed Medical Journals Published in Former Yugoslav Countries. Med Arch. 2020;74(4):252-264.
9. Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, Ulmer H. Statistical errors in medical research-a review of common pitfalls. Swiss medical weekly. 2007;137(0304).
10. Lipman TO. Critical reading and critical thinking--study design and methodology: a personal approach on how to read the clinical literature. Nutr Clin Pract. 2013;28(2):158-64.
11. Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. Kidney Int. 2008;73(2):148-53.
12. Kaczor M, Wójcik R, Połowinczak-Przybyłek J, Potemski P. Critical appraisal of clinical trials in oncology — part I. Oncology in Clinical Practice. 2019;15(2):89–103.
13. Sedgwick PM. Multiple significance tests: the Bonferroni correction. BMJ 2012; 344:e509-e509.
14. Thissen D, Steinberg L, Kuang D. Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False Positive Rate in Multiple Comparisons. Journal of Educational and Behavioral Statistics, 2002; 27(1): 77-83.
15. Sedgwick P. External and internal validity in clinical trials. Bmj. 2012;344:e1004.

16. Bornhöft G, Maxion-Bergemann S, Wolf U et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. BMC Med Res Methodol 2006; 6, 56.
17. Rogozińska E, Gargon E, Olmedo-Requena R, Asour A, Cooper NAM, Vale CL, Van't Hooft J. Methods used to assess outcome consistency in clinical studies: A literature-based evaluation. PLoS One. 2020;15(7):e0235485.
18. Frumento P, Gimelli A. Understanding statistical analysis in randomized trials: tips and tricks for effective review. Eur Heart J Imaging Methods Pract. 2025; 3(1): qyaf036.
19. Austin PC, White IR, Lee DS, van Buuren S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. Can J Cardiol. 2021;37(9):1322-1331.
20. Tenny S, Hoffman MR. Odds Ratio. [Updated 2023 May 22]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK431098/
21. Baguley T., Standardized or simple effect size: what should be reported? Br J Psychol. 2009;100(Pt 3):603-17.
22. Darling, HS. Statistical errors in scientific research: A narrative review. Cancer Research, Statistics and Treatment 7(2):p 241-249, 2024.
23. Sica GT. Bias in research studies. Radiology. 2006;238(3):780-9.
24. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.
25. Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, Tunn R, Aggarwal R, Berkwits M, Berlin JA, Bhandari N, Butcher NJ, Campbell MK, Chidebe RCW, Elbourne D, Farmer A, Fergusson DA, Golub RM, Goodman SN, Hoffmann TC, Ioannidis JPA, Kahan BC, Knowles RL, Lamb SE, Lewis S, Loder E, Offringa M, Ravaud P, Richards DP, Rockhold FW, Schriger DL, Siegfried NL, Staniszewska S, Taylor RS, Thabane L, Torgerson D,

Vohra S, White IR, Boutron I. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ. 2025;389:e081123.
26. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. PLoS Med. 2021;18(3):e1003583.
27. Man-Son-Hing M, Laupacis A, O'Rourke K, Molnar FJ, Mahon J, Chan KB, Wells G. Determination of the clinical importance of study results. J Gen Intern Med. 2002;17(6):469-76.
28. Kara-Junior N. How to measure the precision of clinical study results. Rev Bras Oftalmol. 2014;73(3):133-4.
29. Guerra-Farfan E, Garcia-Sanchez Y, Jornet-Gibert M, Nuñez JH, Balaguer-Castro M, Madden K. Clinical practice guidelines: The good, the bad, and the ugly. Injury. 2023;54 Suppl 3:S26-S29.
30. Field MJ, Lohr KN. Clinical practice guidelines. Directions for a new program. 1990;1990.
31. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Graham R, Mancher M, Miller Wolman D, et al., editors. Clinical Practice Guidelines We Can Trust. Washington (DC): National Academies Press (US); 2011. Appendix C, Clinical Practice Guideline Appraisal Tools. Available from: https://www.ncbi.nlm.nih.gov/books/NBK209533/
32. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.
33. Jaikumar V. Non-Inferiority Trials: Understanding the Concepts - Tutorials and Fundamentals [Internet]. Students 4 Best Evidence. 2022 [cited 2024 Dec 31]. Available from: https://s4be.cochrane.org/blog/2022/03/18/understanding-non-inferiority-trials/